Erasmus University Rotterdam
Erasmus School of Economics (ESE)

Thesis

To obtain the academic degree of
Master of Science in Economics & Business
(Major in Marketing)

Online Personalization in the Age of Big Data

A Trade-off between Value and Concerns for Consumers

Author: Daan van Pinxteren

Email address: d.h.vanpinxteren@students.eur.nl

Supervisor: G.J. Prevo MSc

Study Program: Business Economics

Specialization: Marketing

Date: 22-05-2019 – Final Version

Abstract

This research updates the personalization privacy trade-off for consumers to the era of Big Data, where personalization is becoming more sophisticated and concerns for manipulation and discrimination supplement concerns for privacy. These two concerns have until recently only been described on a theoretical level, but this research operationalizes them for empirical research. It constructed valid scales for the concepts and used those for subsequent hypothesis testing. From a consumer viewpoint, this research assessed the effect of value of personalization and the three concerns on the likelihood of using personalized offerings. It also incorporated a moderating effect of perceived information control. This research once more confirmed the existing trade-off between privacy concerns and value of personalization and also found some evidence that manipulation concerns can lead to less likelihood of using personalized offerings when perceived information control is low. No evidence was found for an effect of concerns for discrimination. Finally, opposite to as was hypothesized, this research found that high perceived control can actually lead to decrease in intent to use personalized offerings when consumers have a high concern for their privacy.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Problem Statement and Research Questions

In the age of Big Data, companies have the ability to acquire massive amounts of personal information from their customers. This information can be used to gain insights into preferences and actions of customers and can consequently be acted upon. For example, companies can make use of personal advertising to target specific customers, leading to an increased effectiveness of those advertisements (Pavlou and Stewart, 2000; Howard and Kerin, 2004). The use of personal information is not only beneficial to companies, but also to consumers themselves, as they receive tailored advertisements that fit their preferences (Chellappa and Sin, 2005). However, over the years, consumers have also expressed concerns of the way companies use their personal data, predominantly because such actions may violate their privacy (TrustE, 2015; Dupre, 2015). This trade-off between privacy and personalization is sometimes coined as the personalization privacy paradox: on the one hand, consumers want more personalized content while on the other hand they care about their privacy that is possibly violated by that content (Awad and Krishnan, 2006). The trade-off between privacy and personalization has been a hot topic in (marketing) literature, with several researchers contributing to it (see e.g. Chellappa and Sin, 2005; Yu and Cude, 2009; Tucker, 2014; Garcia-Rivadulla, 2016).

However, nowadays companies can make use of personal data in an enhanced setting, namely by using sophisticated statistical techniques. Analysing vast amounts of data, companies can derive patterns of consumer behaviour and predict how certain consumers behave in certain settings (Shaw et al., 2001; Calo, 2013). Companies can improve their personalization systems by incorporating consumer behaviour and offer more sophisticated personalized offerings not only based on historical evidence but also possibly based on predictions (Wedel and Kannan, 2016), which makes these companies very powerful (de Rek, 2017). At the same time, concerns for consumers are also becoming broader than just issues of privacy. Several scholars notice other harms of Big Data in a commercial setting, such as manipulation, consumers being manipulated to make certain decisions, (see e.g. Calo, 2013; Kaptein and Eckles, 2010) and discrimination, consumers being treated differently on the basis of Big Data algorithms (see e.g. Hirsch, 2015; Barocas & Selbst, 2016). In the age of Big Data, the personalization privacy trade-off is therefore no longer the full side of the story, meaning that existing research on the topic is slowly becoming obsolete. Therefore, it would be valuable to update the trade-off consumers make when faced with personalization to the modern era of

Big Data by incorporating these new concerns of discrimination and manipulation in the trade-off. This new trade-off is coined in this research as the personalization concerns trade-off.

This research studies the above described trade-off in a commercial setting and from a viewpoint of the consumer. This research aims to answer whether consumers are concerned about harms that may come to them when faced with personalization based on Big Data, which of these harms are most concerning to them and whether consumers have a negative intention towards making use of personalized offerings when they also value these offerings. In other words, this research tries to answer what aspects of the personalization concerns trade-off are relevant to the consumer and which part of the trade-off is favoured: the value of personalization or the harms that may come from it. In addition, to extend the framework, this research incorporates a moderating effect of perceived information control to assess whether consumers are less concerned about the aforementioned harms when they perceive they are in control over what personal information is given.

1.2 Scientific and Practical Relevance

This research aims to build on existing literature on the personalization privacy trade-off and tries to update the trade-off to the age of Big Data, where personalization is becoming more sophisticated and where not only privacy, but also issues of manipulation and discrimination may concern consumers. These additive issues have been addressed in academic papers before (e.g. Calo, 2013; Hirsch, 2015 and Barocas & Selbst, 2016), but only on a theoretical and descriptive level. To the best of my knowledge this is the first study that empirically tests the impact of manipulation and discrimination concerns in a context of online personalization. Furthermore, with more sophisticated personalization due to big data analytics, the value of personalization for consumers may also be higher. This research also tries to incorporate this sophisticated personalization in the trade-off. Therefore, the personalization privacy trade-off will be enhanced and literature on the topic will be supplemented and expanded by incorporating both new concerns and new benefits in the trade-off.

With respect to practical relevance, this research is twofold. First of all, it provides managers a better understanding of customer concerns related to personalization in contemporary times. This may help managers address these issues and implement online personalized advertising and offerings in such a way that consumers respond positively to the advertisements, in a sense that it leads to an increase in performance. Second, this research may give insights in the field of regulation. Policymakers can use this research to determine if and how personalization in the age of Big Data should be regulated if it leads to consumer concerns, in order to protect these consumers in a transactional sphere.

1.3 Structure of the Research

The structure of this research is as follows. First of all, in chapter 2, the theory behind the research is set out on the basis of a review of existing important literature on the topic. First, some general context and clarification of the topic is given after which the conceptual framework used in this research is drawn up. On the basis of this conceptual framework, several hypotheses are proposed and substantiated with relevant academic papers. Then, in chapter 3, the methodology of this research is explicated. The research design is mentioned, after which the concepts are further conceptualized and operationalized into measures. Items are constructed on the basis of existing scales, and for the concerns for manipulation and discrimination, new scales are constructed. Preliminary interviews are held to assess content validity of the scales, after which a final survey instrument is created. In the following chapter 4, actual data analysis is described and results are given. Starting with a description of the sample, then reliability analysis is conducted and construct validity is assessed, after which hypotheses are tested by means of a linear regression. This research ends with a conclusion where the most important results are discussed, scientific and practical implications are given and limitations and directions for future research are provided.

2 Theory and Hypotheses

In this section, the theory on which this research is based is discussed. First of all, the research topic is clarified more, after which an overview and explanation of the conceptual framework is given. The section then continues with the theoretical framework concerning the constructs, starting from the dependent variable. Appendix 1 provides an overview of selected literature related to the topic.

2.1 Personalization in the Age of Big Data

In order to get a better understanding of the research, first the research topic, personalization and concerns for the consumer in the age of Big Data, is clarified further. Both sides of the personalization concerns trade-off are discussed hereunder. Since this trade-off is made in the context of Big Data, first this term is explained.

2.1.1 Big Data

Nowadays, the term Big Data is widespread and widely used. Although there is no clear definition, Big Data is often related to the aspect of a lack of structure within stored data. More specifically, it is related to enormous amounts of data that are no longer identifiable and quantifiable (Cukier & Mayer-Schönberger, 2013). However, volume is not the only aspect of Big Data. The term is also related to the wide variety in data that is stored and analysed. This is mainly due to datafication: everyday aspects of life are quantified, such as locations, but also friendships, romantic relationships (Cukier & Mayer-Schönberger, 2013) and data gathered from the Internet of Things: devices that have an intelligent component and can communicate with other devices for certain goals (Atzori, Iera, Morabito, 2010). Also, Big Data can be described by the velocity at which data is generated and analysed. Users of social media and visitors of websites generate enormous amounts of data which are stored by firms instantly, and which are analysed almost in real-time (Russom, 2011).

The characteristics described above: volume, variety and velocity are also called the 3V's of Big Data and provide a clear overview of what Big Data entails (Gartner IT glossary, n.d.). In a commercial setting, Big Data analysis is nowadays often used for the purpose of sophisticated online personalization in a commercial setting.

2.1.2 Online Personalization

Online personalization is a topic that has been researched extensively over the past years. In a commercial setting, personalization is linked to several topics such as customer trust and loyalty, recommendation systems, comparison agents, privacy concerns and marketing strategy (Adolphs & Winkelmann, 2010). Therefore, there are many different interpretations of personalization and no clear definition of the concept can be given. However, online personalization can be described as tailoring and recommending products and services to consumers on the basis of specific characteristics (Hun Lee & Cranage, 2010). More specifically, it can be described as delivering “the right content to the right person in the right format at the right time” (Ho & Tam, 2005, p.96). Or likewise: “the ability to provide content and services that are tailored to individuals based on knowledge about their preferences and behaviours” (Hagen, 1999 in Adomavicius & Tuzhilin, 2005, p. 83).

Personalization makes use of personal data provided by customers. This data can be analysed to uncover patterns, and ultimately comprehensible information, that can be used for subsequent application, which is called datamining (Shaw et al., 2001). By implementing datamining together with Customer Relationship Management (CRM) tools, companies can create customer profiles. These profiles enable companies to target individual customers (Chellappa & Sin, 2005). For example, customers can receive tailored offerings on websites, individual discounts via email, or advertisements on social media of websites they have visited before (retargeting). This helps companies with improving customer satisfaction, developing customer loyalty and cross-selling opportunities and therefore is a valuable tool (Chellappa & Sin, 2005; Alba et al. 1997; Peppers, Rogers & Dorf, 1999; Adomavicius & Tuzhilin, 2005).

In the era of Big Data, datamining becomes more sophisticated, because patterns can be derived from large amounts of different sources of data which are updated real time. This means in turn that companies can implement personalization in such a way that the individual customer is targeted more precisely because of more accurate and dynamic customer profiles, also called adaptive personalization (Wedel & Kannan, 2016), which in turn leads to more value for companies (Erevelles, Fukawa & Swayne, 2016). Personalization even goes as far as being predictive, in the sense that companies can predict how customers will behave in the future based on Big Data analysis, after which the consumer is steered to desired behaviour for the company (Kaptein & Eckles, 2010).

In this research, the focus is on online personalization in the age of Big Data as part of the personalization concerns trade-off. As is mentioned, this entails personalization in a commercial setting. In other words: personalization that is related to the sales of products and services. Furthermore, since this trade-off is made by individuals, personalization and related terms are further discussed from an individual viewpoint.

2.1.3 **Concerns for the Consumer**

Logically, on the other side of the personalization concerns trade-off are concerns for the consumer that may prevent her from accepting personalized offerings in an online setting. All these concerns stem from the same fact that a consumer has to give up her personal information to receive some form of personalization. As is mentioned in the introduction, most predominantly researched are consumer concerns regarding privacy, with some studies also taking into account the new reality of Big Data (see e.g. Erevelles, Fukawa & Swayne 2016 and Mai, 2016).

However, nowadays, issues regarding privacy are not the only concerns for the consumer in the context of Big Data personalization. Over the past years, several scholars also mention other concerns related to this topic. Nonetheless, these concerns are not so clearly defined as privacy concern, as the studies in which they occur are non-empirical and mostly abstract. However, based on a previous literature review which has been conducted for my master thesis in Law, two different sets of concerns for individuals, next to the concerns for privacy, have been observed.

The first set is related to the impediment of the online freedom of choice of a consumer. By analysing Big Data with sophisticated statistical techniques, companies have the possibility to target their customers when they are not making fully rational choices (Calo, 2013). Furthermore, companies can personalize offerings in such a way that they exploit the vulnerability of an individual consumer (Newman, 2014; Calo, 2013). In other words, consumers can be manipulated by companies to make certain decisions, which may impede their online freedom of choice. Therefore, concerns regarding this phenomenon are coined manipulation concerns.

The second set of concerns is related to the violation of the right to equal treatment. By using Big Data analysis, companies can distinguish personalized offerings between groups of people, such as ethnic groups or age groups (Hirsch, 2015). Likewise, companies can make use of personalised pricing on the basis of Big Data, since they can assess what price a certain consumer is willing to pay based on personal information and online behaviour (Odlyzko, 2003). Consumers may take offense in these actions, especially when it excludes them from certain offerings or if they are charged higher prices. Therefore, concerns about these issues are coined discrimination concerns.

Since concerns for manipulation and discrimination are discussed from an individual viewpoint, this means that no actual manipulation or discrimination has to occur. Neither does it mean that certain laws have to be violated. For example, price discrimination is not forbidden by law, but a consumer may still find it offensive and may perceive the company negatively as a consequence.

2.2 **Conceptual Framework**

Now that the research topic has been explained further, in the next section the conceptual framework for this research is discussed. The conceptual framework is given in the figure below:

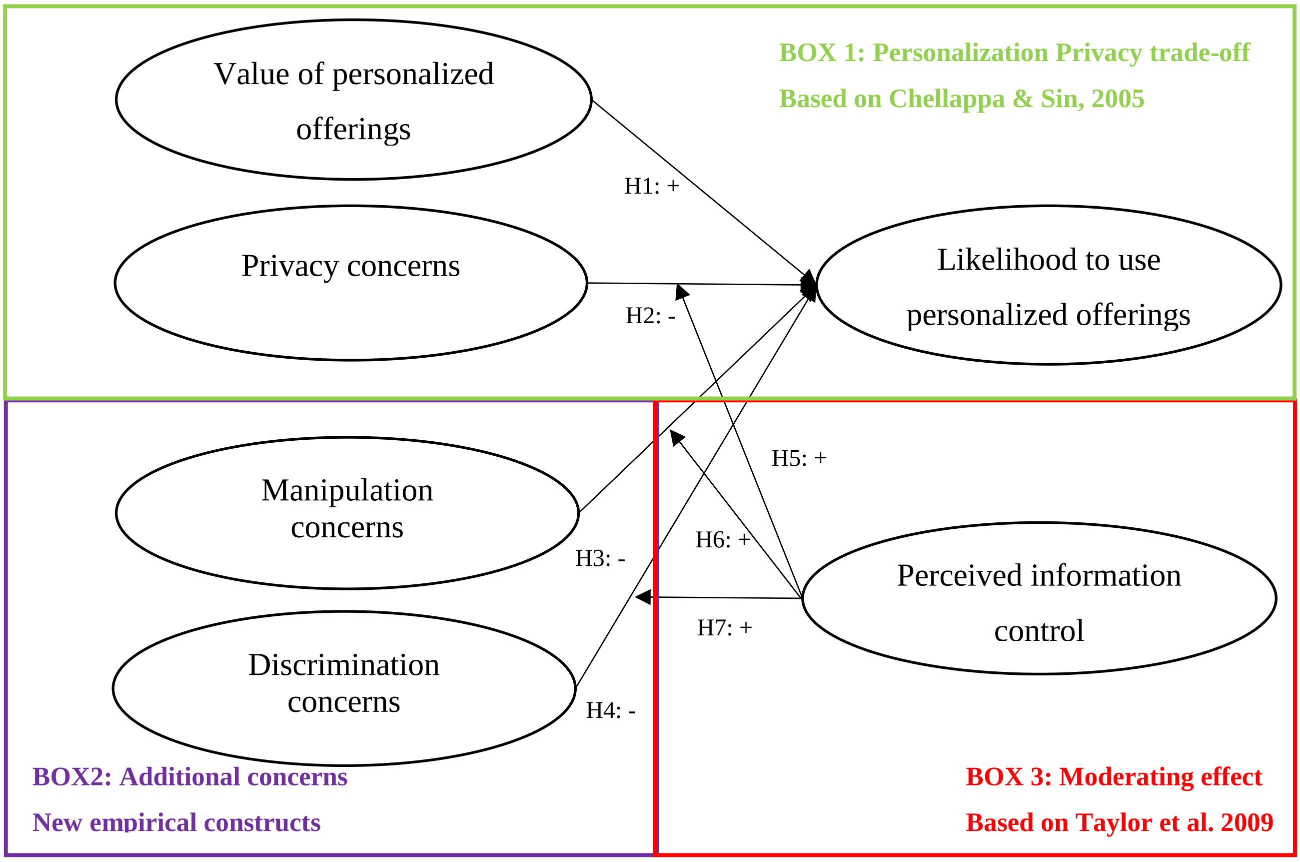

Figure 1. Conceptual framework

As can be seen from the boxes above, the conceptual framework is composed of and inspired by existing research. The constructs value of personalized offerings, privacy concerns and the likelihood to use personalized offerings (DV) are roughly based on Chellappa & Sin (2005) who use a similar framework compared to the one described above. The constructs discrimination concerns and manipulation concerns however are to the best of my knowledge, new constructs which have not been tested in empirical research before. These constructs add to the construct privacy concerns and together they form the concerns side of the personalization concerns trade-off.

The construct perceived information control as a moderator is based on previous work of Taylor, Davis & Jillapalli (2009) which studied the moderating effects of information control and compensation on privacy concerns. The construct information control is expanded in this framework to not only moderate the effect of privacy concerns on the likelihood of using personalized offerings, but also the effect of manipulation concerns and discrimination concerns on the same dependent variable.

Below, the framework is explained more in detail, corresponding to the three boxes that make up the framework. The overview of literature that served as basis for the framework can be found in appendix I.

2.3 **The Personalization Privacy Trade-off (BOX 1)**

Over the past years, privacy related to personalization in the digital environment has been studied extensively. From the 1970s onwards, there exists a notion of a so-called calculus of behaviour when individuals make decisions involving their personal information (Laufer and Wolf, 1977). In this calculus, individuals take into account anticipated benefits and unpredictable consequences, which serve as a predictor whether individuals would disclose personal information, accounting for situational constraints such as institutional norms. Based on Laufer and Wolf (1977), Culnan and Armstrong (1999) argue that, in the specific context of buying products and services, consumers assess the likelihood that their information will be used fairly before they decide to give their personal information. More specifically, when consumers are informed about the information practices of the retailer and perceive the business as fair, they are more willing to disclose personal information. In subsequent literature, this privacy calculus has been linked to a cost-benefit analysis, consistent with expectancy theory that states that individuals behave in ways that maximize positive outcomes and minimize the negative (Culnan and Bies, 2003; Dinev and Hart, 2006). In other words, when individuals give up their personal information, they assess whether the overall benefits of disclosure are balanced by or less than the costs of disclosure, for example the perceived risk involved.

The privacy calculus and the respective trade-off between benefits and costs has been empirically tested, mainly in an online context. Dinev & Hart (2006) study the effect of

Internet Privacy Risk (the costs) and Personal Internet interests (the benefits) on the willingness to provide personal information on the Internet. The authors found that both factors significantly influence the willingness to provide personal information on the internet. In the more specific context of online personalization, important for this research, Chellappa and Sin (2005) studied the effect of privacy concerns and value of personalization on the likelihood of using personalized services.

These authors also found that both costs and benefits influence a consumer's decision to disclose personal information. Thus, it can be said that on the internet, consumers value personalization on the one hand, but care about their privacy on the other hand. Awad & Krishnan (2006) coin this phenomenon as the personalization privacy paradox: consumers value online personalization but want to give up as little personal information to receive this personalization. This is paradoxical in the sense that personal information is needed to actually offer personalization to the consumer (Xu, Luo, Carroll & Rosson, 2011). There exists a tension between commercial companies' exploitation of consumer information to offer personalization, and those consumers' concerns about the privacy of that information (Sutanto, Palme, Tan & Phang, 2013).

In subsequent research, the privacy calculus or privacy paradox has also been studied in the context of more modern, specific forms of personalization. Sheng, Nah & Siau (2008) study the effects of privacy concerns and perceived benefits on the likelihood to adopt u-commerce (ubiquitous commerce; commerce that is targeted at anyone, anywhere at any time). Furthermore, Xu, et al. (2011) study the effects of the same constructs on the willingness to have personal information used in location aware marketing.

In this research, the above described notion of the privacy calculus is applied on personalization in the age of Big Data. Following the conceptualization in Chellappa & Sin (2005), this research studies both the value of personalization and privacy concerns on behavioural intentions related to Big Data personalization (the likelihood to use personalized offerings). These constructs are hereunder explained further.

2.3.1 The Likelihood to use Personalized Offerings (DV)

In previous research related to the personalization privacy trade-off, the dependent variable has taken many forms. As discussed above, Chellappa and Sin (2005), conceptualize the outcome variable as the likelihood of using personalized services. Sheng, Nah & Siau (2008) and Awad & Krishnan (2006) use a similar construct, where the latter make a distinction between personalized services and personalized advertising. Furthermore, Taylor et al. (2009) conceptualize the variable as behavioural intentions linked to personalization. Then, multiple scholars such as Dinev & Hart (2006), John, Acquisti & Loewenstein (2010) and Jai & King (2016) conceptualize the variable as the willingness to provide personal information for personalization purposes. Hun Lee & Cranage (2010) use both the willingness to give up personal information and the likelihood to adopt personalized services as a dependent variable.

As discussed above various dependent variables have been used in previous studies. However, these different conceptualizations have some degree of similarity: each construct is linked to the perception of the consumer towards personalization and each construct focuses on behavioural intentions, be it by explicitly researching whether a consumer is likely to use a personalized offering, or more indirect, by focusing on the intention to give up personal information for personalized offerings. The similarity between the constructs is evident from Chellappa & Sin (2005), who use the willingness to provide personal information as a proxy for the likelihood.

In this research, based on Chellappa & Sin (2005), the dependent variable is the stated likelihood that consumers will use personalized offerings which are provided by companies based on Big Data analysis, because this likelihood most directly reflects what the research aims to answer: how consumers perceive personalization in the age of Big Data. However, it should be mentioned that this likelihood to use personalized offerings is based on consumer rationale, and not on decisions made in the heat of the moment. Thus, the variable, like other conceptualizations, is closely linked to behavioural intentions of a consumer (Fishbein & Azjen, 1975; Taylor et al. 2009), namely: the intention of a consumer to make use of a personalized offering. However, within the framework of reasoned action, it can be argued that behavioural intention approaches actual behaviour (Fishbein & Azjen, 1975).

In this research, personalized offerings are defined as: personalization that is centred on the purchase of products and services. This may include personalization of (aspects of) products and services themselves such as personalized product features, product design or service packages, but also personalization of the purchasing process itself such as recommendation systems, personalized prices or discounts, and personalized advertisements. In the latter case, the products and services themselves are not differentiated, but the process for purchasing these products is (see e.g. Kaptein & Eckles, 2010).

2.3.2 The Value of Personalized Offerings

As discussed before, personalization is beneficial to companies, as it allows them to target customers individually, which can lead to increased customer satisfaction, retention and cross-selling. However, from a viewpoint from individuals, personalization can also be beneficial. Personalization leads to increased convenience, time saving and individualization which are important objectives for customer value (Sheng, Nah & Siau, 2008). Furthermore, personalization can lead to decreases in information overload which can lead to increased customer satisfaction (Li & Unger, 2012; Sutanto et al. 2013; See also West et al. 1999). Therefore, it can be said that personalization has some value to consumers. Chellappa and Sin (2005) argue that the value of online personalization for a consumer mainly stems from the fit that a personalized offering, namely a product or service or related aspects such as an advertisement, provides. In addition, they argue that online personalization is valuable because personalized offerings are delivered in a proactive fashion, for which the consumer's decision making is simplified. The net benefit of online personalization can be seen as the convenience of having parts of the online browsing and purchase experience personalized. It can further be argued that the more quality personalization has, the more value this gives (Li & Unger, 2012). Therefore, personalization in the age of Big Data, which is highly sophisticated, can be said to be even more valuable to consumers.

Chellappa & Sin (2005) found that perceived value of personalization is an important factor with regard to stated intentions to use personalization services. Furthermore, they discovered that the value of personalization outweighed the privacy concerns. This may be due to the notion that giving up personal information is seen by the consumer as a necessary monetary transaction to get benefits from that personalization (Schumann, Wagenheim and Groene, 2014). White (2004) confirmed that users are more likely to provide personal information when they receive benefits of personalization. Consumer surveys indicate that the majority of consumers value online personalization and are willing to give up personal data to receive it (Kobsa, 2007). However, not all consumers value personalization equally, since some consumers may have different online preferences with respect to personalized offerings, dependent on their personality and online behaviour. Thus, it can be argued that the value of personalization is consumer specific. In this research, the value of personalization is conceptualized as the value of personalized offerings. The more value a consumer has for personalized offerings, the more likely she is to make use of those offerings by a company.

H1: The likelihood to use personalized offerings is positively influenced by the perceived individual value of these personalized offerings.

2.3.3 **Privacy Concerns**

Privacy in a context of online personalization is more specifically information privacy. This refers to "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" (Westin, 1967, p. 7). Information about individuals is mainly information that can identify individuals, so-called personally identifiable information (FTC, 2000) such as email addresses, names and phone numbers. However, an aspect of Big Data analysis is that it also makes use of information that cannot identify a specific person, so-called personally unidentifiable information such as age, date of birth and gender and even anonymous information such as IP address and domain type. Customer profiles are created based on pools of this unidentifiable or anonymous information, after which these profiles can be merged with personal information. This way, profiles that are created with these types of data can later be used to identify specific individuals and target them (FTC 2009, FTC 2012), also called re-identification (Tene & Polonetsky, 2013). Thus, personally unidentifiable and anonymous information can be linked to specific consumers as well. This means that collection of these types of information could also be marked as collection of personal information and thus be of concern to the consumer (Chellappa & Sin, 2005. See also the General Data Protection Regulation, Regulation (EU) 2016/679 (hereafter: GDPR), article 4).

Privacy concerns therefore are related to the violations of informational privacy, including personally identifiable, personally unidentifiable and anonymous information. However, as the focus of this research is on individual concerns for privacy, it means that privacy does not actually have to be violated by a company in order for it to be a concern for the consumer. As long as the consumer perceives that her privacy is violated, it may result in a different view towards the company. This is also called subjective privacy harm (Calo, 2011; Taylor et al. 2009). Thus, privacy concerns refers to an individual's subjective view of fairness within the context of information privacy (Campbell, 1997).

Smith, Milberg & Burke (1996) argued that consumers' concerns for privacy consist of the following four dimensions: (1) collection, which reflects the concern that large amounts of personal information is stored and collected, (2) unauthorized secondary use, which is the concern that information is collected for one purpose, but used for something else, (3) improper access, the concern that information is available to people who are not authorized by the individual to use it and (4) data errors, which reflects the fear that there is no adequate protection between deliberate and accidental errors such as data breaches. In other words, concerns may entail consumers' worry that personal information is used for unwanted targeted advertising, sold to other companies, reviewed by unauthorized individuals, or even hacked, but also simply that personal information is gathered (Li & Unger, 2012). Nonetheless, in the age of Big Data, it is impossible for individuals to have full control over their informational privacy when engaging on the internet, as information about an internet user is gathered everywhere. Thus, the concern that personal information is collected is becoming less relevant. More important is that streams of individual information move ethically (Richards & King, 2014). Therefore, concerns that personal information may be misused become more prevalent.

For simplification reasons, it is argued in this research that privacy concerns only entail concerns about personal information that is stored accurately by a company and is protected from improper access. Thus, concerns about improper access and data errors are not within the scope of this research, as this leads away from the focus on online personalization. Privacy concerns therefore are made up of concerns of collection of personal information and unauthorized secondary use.

To obtain personalized offerings, consumers need to provide personal information about them so that the company can tailor its offerings to their preferences. Therefore, it is unavoidable that the consumer loses some of her informational privacy in order to make use of effective personalized offerings. This means that a consumer who is concerned about her privacy, may ultimately choose not to be involved in these personalized offerings. This is confirmed in many studies, where consumer concerns for privacy are negatively related to behavioural intentions with regard to personalized offerings (Malhotra, Kim & Agarwal, 2004; Chellappa & Sin, 2005; Awad & Krishnan, 2006; Dinev & Hart, 2006; see for a complete overview Mothersbaugh, Fox, Beatty & Wang, 2012). In addition, multiple consumer surveys also acknowledge privacy concerns as a reason for consumers not to disclose personal data to a website (Kobsa, 2007).

On the other hand, it can be stated that individual concerns for privacy are low, since most information is gathered by companies with consent from the online consumer, as is often required by law (see e.g. GDPR, article 6 and 7). Should consumers be concerned about their privacy, they would not give their consent, and the fact that many online consumers do consent to give personal information may mean that they are not concerned about their privacy. However, Dinev & Hart (2006) found that giving consent may still mean that consumers are concerned about their privacy, but that they take other factors into account to favour the decision to disclose personal information (the cost-benefit analysis of the privacy calculus). Another reason for consumers to disclose personal information may be that consent is easily given without much rationale, for example because privacy policies and other aspects that require consent are (deliberately) too complex and extensive for the consumer (World Economic Forum Report, 2013). It may well be that a consumer has a concern for her privacy but makes a different decision in the heat of the moment. All in all, the simple fact that consumers disclose personal information with consent does not necessarily mean that they have low privacy concerns.

Therefore, it can be said that privacy concerns have a negative influence on the likelihood to use personalized offerings. But, also concerns for informational privacy are individual specific (Malhotra, Kim & Agarwal, 2004; Smith, Dinev & Xu, 2011). Consumers can be categorized by their concern levels, from individuals who are highly concerned about their rights to the unconcerned. In between are pragmatists, whose privacy concerns are moderate and who represent the largest portion of consumers (Taylor, 2003 in Alhouti, Johnson & D’Souza, 2016). Therefore, it is argued that the likelihood to use personalized offerings is related to the individual specific concerns for their information privacy.

> **H2:** The likelihood to use personalized offerings is negatively influenced by the individual’s concerns for online privacy.

2.4 **Additional Concerns (BOX2)**

Together with individual’s concerns for privacy, the new empirical constructs concerns for manipulation and concerns for discrimination are incorporated in this research. These concerns complement the concerns for privacy and together they form the negative side of the personalization concerns trade-off. It has to be noted that these concerns stem from the same information gathering of the user. Therefore, it may be likely that a consumer who is concerned about her informational privacy, will also be concerned about manipulation or discrimination.

2.4.1 **Manipulation Concerns**

As sophistication of personalized offerings in the age of Big Data grows, consumers may be exposed to a new issue regarding their online behaviour, namely the impediment of their freedom of choice (see e.g. Calo 2013; Helberger 2017). By making use of advanced profiling techniques involving machine learning aspects such as decision trees, personalization becomes predictive; companies can extract how individuals will behave in future situations. Consequently, these predictions can be acted upon by companies (Calo 2013; Angwin 2014. See also Bilenko & Richardson, 2011). More concretely, companies can formulate personalization in such a way that individual behaviour is steered to maximize effectiveness of the offering. A company can model how a certain consumer would behave rationally in a certain situation. Then, by analysing personal information such as purchase history, it can assess when and how a certain consumer does not behave rationally. These customers can be targeted at their most vulnerable moments, such as after a hard day's work (Calo, 2013). Likewise, computer algorithms on the basis of Big Data can set prices for each individual that correspond with her maximum willingness to pay (Angwin, 2014).
Moreover, influence strategies can be specified per customer profile. Kaptein & Eckles (2010) coin this phenomenon as persuasion profiling. Profiles of consumers are created that capture variation in responses to influence strategies. An individual's persuasion profile indicates which influence strategies are expected to be most effective (Kaptein & Eckles, 2010). This may lead to an advertisement as being customized to 'only one left in stock' for a customer that is vulnerable to scarcity, whereas another customer vulnerable to authority may see the same product but then advertised as 'professor X. recommends the product'. By making use of personal information, companies can manipulate consumers to make certain decisions.

Online manipulation can lead to the exploitation of vulnerable consumer groups, such as elderly people. For example, people with Alzheimer's could be identified and offered the same product twice (Newman, 2014; Angwin 2014). In practice, Facebook has been accused of targeting young people when they are feeling stressed or anxious, and thus use their emotional state to boost response to their advertising content (Hutchinson, 2017). Therefore, online manipulation can lead to negative consequences for these groups of consumers. However, a property of Big Data is that weaknesses of all consumers, and not only the vulnerable groups, are exposed. The differences between vulnerable consumer groups and the average consumer are becoming more blurred (Calo, 2013). Therefore, it can be argued that every consumer could be the target of online manipulation and steered towards certain behaviour, beneficial for the company.

As a result of this, the consumer's freedom of choice is impeded. Thus, consumer concerns for manipulation can be defined as the degree to which consumers find that companies impede their online freedom of choice with personalized offerings. Again, these concerns are individual specific. Some consumers may attach more value to their freedom of choice and the possibility to 'be in control' than others. Analogously to privacy concerns, it can be said that when a consumer has concerns for manipulation, she will steer away from personalized offerings by companies, since these offerings are the outcomes of this online manipulation that steer consumers to a certain online behaviour and impede their freedom of choice. Therefore, the following hypothesis can be offered:

> **H3:** The likelihood to use personalized offerings is negatively influenced by the individual's concerns for being manipulated towards certain behaviour online.

2.4.2 **Discrimination Concerns**

In order for online manipulation of the consumer to work, personalized offerings have to be differentiated between certain consumers, which in turn entails that consumers are prone to online discrimination. However, the simple fact that consumers are put in certain profiles does not necessarily lead to problems of inequality. Furthermore, the fact that consumers receive different offerings may also not be problematic, since this is inherent to personalization. Only when a consumer is negatively affected compared to other consumers when having a certain profile, for example when she is excluded from discounts on the basis of race, discrimination becomes noticeable and problematic to that consumer (see e.g. Schrage, 2014). For example, Steel & Angwin (2010) found such discrimination by Capital One, an American credit card company, that showed different credit cards with different rates to different website visitors, based on their customer profiles involving geographic area an income (see also Angwin, 2014). Furthermore, Sweeney (2013) has found such discrimination in advertisement delivery on the Google search result page (Google Ads) that was based on racially associated names, including advertisements about criminal arrest records. These examples, of course, could be problematic to those consumers involved.

Discrimination can occur if certain information that could in its core be used discriminatively, such a race and gender, are intentionally captured as variables in the Big Data analysis that forms the basis for personalized offerings (Barocas, 2014). However, discrimination can also occur more indirectly, namely when information that could be used to discriminate individuals is not deliberately added in the analysis, but correlates with other variables, such as geographic areas, and thus is a latent variable in the analysis (Hirsch, 2015; Barocas & Selbst, 2016). A company can therefore be unconscious of the underlying discrimination in the analysis (Barocas, 2014; Barocas & Selbst, 2016; Hirsch, 2015). Therefore, a company should not be deliberately discriminating in order to raise concerns for the consumer.

A specific form of online discrimination with regard to personalization that is becoming widespread is price discrimination, in the sense that specific individuals (with certain customer profiles) receive different prices for the exact same product or service, based on personal information such as website browsing behaviour (Odlyzko, 2003; Shiller, 2014). For example, on the travel website Orbitz, Mac users were guided to pricier hotels as they usually spend more per night (Mattioli, 2012). In addition, Angwin (2014) found that online stores such as Staples.com, charged more for people who are browsing with zip codes of areas that have fewer rival stores. Hannak et al. 2014 actually found that 9 out of 10 big companies such as Best Buy, HomeDepot and JCPenney used some form of price discrimination among users. Consumers subject to this price discrimination can find this of great concern.

A further problem that rises with regard to online discrimination by making use of Big Data algorithms, is the possibility that once a customer is profiled, she is discriminated on the basis of that profile in several occasions over the course of her life (Citron & Pasquale, 2014; Angwin, 2014; see also Hirsch 2015). Citron & Pasquale (2014) give the following example: "Imagine a young woman who failed to get a job out of college, and that failure reduced her "employability" score used by potential employers to determine her fitness for work. She found part-time work at a fast food restaurant. Her credit score fell far below 600 without her even knowing it, perhaps because of inferences associated with certain low-paying jobs. Her low credit score *caused* further bad outcomes, cascading into ever more challenging life circumstances. Talent analytics companies categorized her as a "non-innovator" and "waste." With low scores across countless measures, the young woman was unable to get a full-time job." (Citron & Pasquale, 2014, p. 33). This repeated discrimination is even more problematic if consumers are mistakenly placed in groups and thereafter discriminated, which could happen since datamining makes determinations about individuals on inferences and correlations and not on facts, and therefore is prone to some bias or incomplete information (Ramirez, 2013 in Barocas, 2014; Bodle, 2014). Thus, it can be argued that all consumers, be it intentional or unintentional, technically correct or incorrect, may be prone to discrimination by commercial companies online.

But, also with regard to online discrimination, it has to be noted that concerns for the consumer are individual specific. Some consumers may attach more value to equal treatment, especially if they have been prone to any form of discrimination in the past. Other consumers likewise, may find discrimination less problematic. Once a consumer has some concerns that she may be discriminated by Big Data analysis used for personalization, it may be likely that she will react negatively towards this personalization. Therefore, the following is hypothesized:

> **H4:** The likelihood to use personalized offerings is negatively influenced by the individual's concerns for being discriminated online.

2.5 Moderating Effects of Perceived Information Control (BOX3)

In previous research, the personalization privacy trade-off has been expanded to capture variables with a moderating or mediating effect, such as perceived user control over information that is extracted online. However, such moderating effects have not yet been empirically tested on the other concerns, which is logical since these concerns form new empirical constructs. In this research, the moderating effect that captures perceived information control regarding personal information is captured in a moderator that is applicable to all three concerns.

2.5.1 Perceived Information Control

In psychology research, it is indicated that perceived control influences people's emotions and behaviours (Averill, 1973 and Skinner, 1996 in Hajli & lin, 2016). Thus, individual perception of control contributes to one's desire for actual behaviour (Hajli & Lin, 2016). In a setting of disclosing personal information, several studies found that perceived control over information, such as knowledge over what information is collected, for what purposes it is used and the possibility to opt out, has a direct effect in influencing the willingness to disclose personal information or the likelihood to make use of personalized offerings (for an overview see Mothersbaugh et al. 2012). For example, Nowak & Phelps (1995) demonstrated that people are less worried about data collection when they are given the choice to opt-out. Brandimarte, Acquisti & Loewenstein (2012) found that increases in perceived control led to increase in the willingness to provide personal information. So do Meinert, Peterson, Crisswell & Crossland (2006), who found that priming with privacy statements on a website that involve a high protection of privacy, lead to more willingness to provide personal information. Similarly, within the context of personalization on social networks, Krasnova et al., (2010) and Hajli & Lin (2016) found that perceived information control increases the desire to share personal information. Furthermore, Tucker (2014) found that perceived control on Facebook over privacy information increases the effectiveness of personalized advertising. Several consumer studies also found that users are more willing to disclose personal data if they possess knowledge of and/or control over the use of this data (Kobsa, 2007).

2.5.2 **Information Control X Privacy Concern**

However, not only does perceived information have a direct effect on the willingness to share information online or to make use of personalized offerings, it can also be argued that, in a model where privacy concerns are incorporated as a direct effect, perceived information control has a moderating effect, in which perceived information control suppresses the relationship between privacy concerns and the likelihood to use personalized offerings.

As mentioned before, a person's informational privacy is "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" (Westin, 1967, p. 7). This definition shows that informational privacy is for a large part dependent on the control of the individual to share that information with others, including companies, and transparency by those companies what is done with the individual's information. Sheehan & Hoy (2000) found that information control is a primary factor for consumers' concerns for privacy online. This is supported by Malhotra et al. (2004), who argue from social contract theory, that collection of individual information is only perceived as fair when the consumer is granted control over this information and she is informed about the company's intended use. They stated that the construct for privacy concerns is a second order construct based on the user's view of collection of information, the user's importance regarding information controls and finally the user's importance that she is aware of the intent of information gathering (Malhotra et al. 2004).

What follows is that, if a consumer attaches much value to information control, she has high concerns for privacy. Then, if a company has a low perceived level of information control, the consumer with high privacy concerns may be likely not to engage in the relationship and stay away from personalized offerings. On the other hand, a high level of perceived information control may decrease the negative effect of privacy concerns on behavioural intentions. This moderation effect of perceived information control has been conceptualized and empirically validated by Taylor et al. (2009), who assess this effect regarding behavioural intentions (namely the use, recommendation and reflections of travel websites). Similarly, Martin, Borah & Palmatier (2017) found that perceived information control suppresses the effect of perceived data vulnerability (including privacy concern) on negative behaviour towards a company (negative word of mouth, switching behaviour and falsifying information). In this research, regarding personalized offerings, therefore the following hypothesis is expected to hold:

> **H5:** The relationship between concerns for privacy and the likelihood of using personalized offerings is suppressed by the perceived control over information

2.5.3 Information Control X Manipulation Concerns & Discrimination Concerns

As mentioned previously, perceived information control has not yet been empirically captured as a moderating effect on the relationship between manipulation concerns or discrimination concerns and the dependent variable. However, it could possibly be captured as a moderator in these relationships as well, in the same way it is hypothesized with regard to the relationship between privacy concerns and the likelihood to use personalized offerings. As research has found that the consumer's perceived control on the collection and use of personal data has a direct effect on the likelihood to use personalized offerings, and, corresponding to privacy concern, manipulation concerns and discrimination concerns originate from that same collection and use of personal information, perceived control may well suppress the relationship between manipulation concerns or discrimination concerns and the likelihood to use personalized offerings as well. A consumer may have a high concern for online manipulation and discrimination, but because she perceives that she can control what personal information is collected and how it is used, she may be willing to make use of the personalized offerings despite having concerns for manipulation and/or discrimination.

This can be clarified further for each set of concerns:

Manipulation concerns

Online consumer manipulation is possible due to the vast amount of information that is gathered from the consumer. For example, advertising with different advertisement headlines for different consumer groups, that tries to persuade these specific consumer groups to buy a product, is possible due to analysis of consumer behaviour and preferences online. The attempt to persuade the consumer may be a concern to her, but if the consumer feels that she has some influence on the manipulation process, by choosing what information she releases to the company, she may still make use of personalized offerings. In other words, she may feel less vulnerable to the manipulation attempt. This is supported by Baker, Gentry and Rittenburg (2005) who argue that perceived control over a situation is a primary factor in the experience of consumer vulnerability in a marketing setting. Consumer vulnerability occurs when the consumer feels that she has no control over the marketing situation (e.g. a persuasion attempt) and consumers feel less vulnerable when they have more control. Aguirre et al. (2015) transpose the ideas of consumer vulnerability to a setting of online personalized advertising. They found that consumers respond less negatively to personalized advertisements when they perceive control over the situation. As personalized advertisements are created on the basis of personal information, this perceived control is predominantly information control. Thus, a company that signals what information is gathered and asks for (implicit) consent for the purpose of highly personalized advertisements, is more likely to succeed in selling than a company who does not.

Therefore, even though a consumer can have a concern that she is manipulated into buying online, she may still choose to accept personalized offerings because she perceives that she has control over what information she releases and therefore thinks she is less vulnerable. Similarly, the consumer may feel more vulnerable when she perceives low control over the situation, which increases the negative effect of manipulation concerns on the likelihood to use personalized offerings. The following is therefore hypothesized:

> **H6:** The relationship between concerns for manipulation and the likelihood of using personalized offerings is suppressed by the perceived control over information.

Discrimination concern

Discrimination, both offline and online, occurs on the basis of personal information. This is logical, as it is personal information such as race or gender that allows discrimination between people. In an online consumer setting, discrimination can occur when certain (groups of) consumers do not receive the same kind of offerings. The consumers who are treated less beneficial in such a situation, for example when they do not receive discounts on products that others receive, may find this type of discrimination a concern. But, if the consumer knows that no information is gathered by the company that could be used to discriminate her, or that she can opt out of collection of information of this type (e.g. not filling in gender in an online registration form), she may respond favourable to the personalized offering. In other words, if the consumer perceives control over what information is collected and used, she may be more likely to use personalized offerings, because she thinks that the company is not likely to discriminate her in such instances, for example because the company knows too little about her to treat her differently. The less control a consumer has, the more she may respond negatively to personalized offerings, since she may feel that personalization is based on information that may be used to discriminate her. It is partly therefore that commissioner Brill of the FTC argued in 2013 that consumers need to be reasserted some control over their information in order for Big Data personalization to be accepted by the consumer (Brill, 2013).

Of course, when a consumer can opt out to provide sensitive information, it may well be that discrimination still occurs, namely due to indirect discrimination, which means that no information that can be used discriminatively is directly collected, but this information correlates with other types of personal information that is collected (see the example in paragraph 2.4.2). However, as long as the consumer perceives that she is less prone to discrimination due to information control, she may still be willing to make use of personalized offerings. Thus, if the consumer has a concern for discrimination, but also has control over the information she provides to companies, which makes her think it is not probable that she is discriminated, she will still make use of personalized offerings by a certain company. On the other hand, the concerns for discrimination will have a stronger negative effect on the likelihood to use personalized offerings if the consumer perceives little control over her personal information, as she is more likely to believe that the personalization has occurred on the basis of discrimination. This leads to the following hypothesis:

> **H7:** The relationship between concerns for discrimination and the likelihood of using personalized offerings is suppressed by the perceived control over information.

3 Method

Now that a literature review has been conducted and the hypotheses have been described, in this next section the method for this research is explained. First of all, the research design is discussed, which is followed by further conceptualization and operationalization of the constructs for the purpose of the research instrument. Here, the measures of the dependent variable, the independent variables and the moderator are described. Important with respect to this part is the creation of measurement scales for the new empirical constructs manipulation concerns and discrimination concerns. After this, results of preliminary interviews assessing face validity of the research instrument are discussed and a description of the final instrument is given.

3.1 **Research Design**

In order to empirically test the hypotheses, this research makes use of a survey. Reasons for choosing a survey over other data collection methods, such as a (survey) experiment are mostly time and cost efficiency, especially with regards to the number of variables in this research. In addition, by making use of a survey, the terms personalized offerings and personal information can be described consistently and in detail, to provide for more accurate results. Finally, a survey makes anonymous data collection possible, which is important for the delicate subject of trading personal information for personalization. The survey is distributed to online consumers, incorporating items for each construct described in the conceptual framework. As such, the survey tries to measure all variables directly. Direct measurement rather than manipulation of certain variables is chosen since all variables measure consumer perception or sentiment (<u>willingness</u> to give information, <u>value</u> of personalized offerings, individual <u>concerns</u> and <u>perceived</u> control), which makes it difficult to manipulate them.

Most of the items used in the survey are extracted from previous research. The same applies for the corresponding scales. However, with respect to the new empirical constructs of manipulation concerns and discrimination concerns, new items and corresponding scales are developed, building on the theoretical framework in the previous section and taking into consideration similar scales, such as the scale for privacy concerns. To improve predictive validity, for all variables multi-item scales are used (Diamantopoulos, Sarstedt, Fuchs, Wilczynski & Kaiser, 2012). In addition, to create an optimal survey instrument, preliminary interviews are held to assess face validity of all items, especially the newly developed ones, before the final survey instrument is created.

3.2 Further Conceptualization and Operationalization

In the following section, the constructs that are described in the theoretical framework are further conceptualized and in turn operationalized as measures. This is explained in detail per variable.

3.2.1 DV: The Likelihood to use Personalized Offerings

As mentioned in the theoretical framework, the dependent variable has taken many forms in research over the years. In this research, the dependent variable is the likelihood that a consumer makes use of personalized offerings by a company. This likelihood is intentional in the sense that it captures the rationale of the consumer. Consumers' intent to make use of personalized offerings is related to their intent to give up personal information for these offerings, as this is necessary to receive such offerings. Therefore, to approximate the likelihood of using personalized offerings, based on Chellappa & Sin (2005), in this research a unidimensional proxy is used to further conceptualize and test the dependent variable, which is the intent to provide personal information for personalized offerings. Another reason for using a proxy is to distinguish the dependent variable from the independent variable 'value of personalized offerings', which may be interpreted quite similarly to 'likelihood to use personalized offerings'. In addition, the willingness to provide personal information serves as a good proxy for whether one actually reveals personal information at the request of a company (Malhotra et al. 2004).

The willingness to provide personal information for personalized offerings has been assessed as a latent variable with several valid measurement scales in research over the years. Since this paper mainly builds on the research of Chellappa & Sin (2005), their measurement scale for the willingness to provide personal information is adopted. This includes a single-item measuring stated comfort to provide personal information using a 7-point Likert scale anchored with strongly disagree and strongly agree (Chellappa & Sin, 2005). However, to provide extra predictive validity, another valid single-item scale is incorporated in the research, which is based on Son & Kim (2008) who in turn base the item on Smith et al. (1996). This item measures the stated likelihood of refusal to receive personalized offerings in return for personal information and is measured using a 7-point Likert scale anchored with strongly disagree and strongly agree. The following scale is therefore used in this research:

Table 1

Scale for likelihood of using personalized offerings (proxy by willingness to provide information)

Item name	Item text	Measurement
LIKPER1	I am comfortable providing personal information to companies in return for personalized offerings	7-point Likert Scale
LIKPER2	I refuse to receive personalized offerings when I have to give personal information in return	7-point Likert Scale

Note. LIKPER2 has been re-coded in analysis as it is negatively worded.

3.2.2 **IV1: The Value of Personalized offerings**

The construct value of personalized offerings, or value of personalization, has first been developed in the paper of Chellappa & Sin (2005). The authors argue that value of personalization consists of the aggregated value of different types of personalization, namely non-purchase related customer attributes, personalization of the product browsing and purchasing experience, such as advertisements, and personalization of products and services themselves (Chellappa & Sin, 2005). More specific, a distinction is made between value of personalized websites and value of personalized products and services. Their scales however are too specifically focused on different types on personalization and are therefore not incorporated in this research, where such distinction between types of personalization is not made.

Other authors argue that the value of personalization can be equated with the attitude towards personalization, which means that attitudinal scales can be used to measure the construct (see e.g. Ho & Kwok, 2003; de Pechpeyrou, 2009). Holbrook & Batra (1987) constructed such scales for consumer attitudes towards advertisements and brands, where attitudes were measured on 7-point semantic differential scales. De Pechpeyrou (2009) transposed this attitudinal scale to a setting of online personalization and measured the items on a 7-point Likert scale. In this research, that scale is adapted, where inspiration for items is also gathered from the researches of Ho & Kwok (2003) and Chellappa & Sin (2005). This resulted in 4 items, all measured on a 7-point Likert scale anchored with strongly disagree and strongly agree. After preliminary interviews however, based on the suggestions, some further adaptions have been made to the scale (see paragraph 3.3). The scale final can be found below.

Table 2

Scale for value of personalized offerings

Item name	Item text	Measurement
VALPER1	I think that personalized offerings suit my needs	7-point Likert Scale
VALPER2*	I like offerings that are personalized for me	7-point Likert Scale
VALPER3*	I have positive feelings towards personalized	7-point Likert Scale
VALPER4	I value offerings that are personalized for my preferences	7-point Likert Scale

Note. *****VALPER2 and VALPER3 were revised after preliminary interviews, see 3.3. See for the original items before revision table 7.

3.2.3 IV2: Privacy Concerns

Information privacy concerns have been conceptualized and operationalized in many studies. The first conceptualization of privacy concerns involved a single dimensional construct, measuring global information privacy concerns (Smith et al. 1996). To understand the complexity of individuals' privacy concerns, Smith et al. (1996) developed a multidimensional scale with 15 items, that reflect four dimensions of information privacy concern, namely collection, unauthorized secondary use, improper access, and errors. The items were measured on a 7-point Likert scale anchored with strongly disagree and strongly agree. Stewart & Segars (2002) empirically confirmed this scale. However, Malhotra et al. (2004) argued that information privacy concerns may in fact be a second-order construct that governs the underlying dimensions (which are then first-order constructs). They conceptualize informational privacy concerns as a second-order construct that is formed by the first order constructs of collection, control and awareness. This led to the development of 10 items, measured on a 7-point Likert scale. Nevertheless, most successive papers that incorporate the construct of information privacy concerns did not find any problems related to the scale constructed by Smith et al. (1996) (Li, 2011). In fact, the scale constructed by Smith et al. (1996) remains the most used in subsequent research and has attained certain empirical reliability, although oftentimes only part of the scale, fitting to a certain research, has been incorporated (Li, 2011).

Also, in this research the scale from Smith et al. (1996) is used. Here, items regarding the collection and unauthorized secondary use dimensions are incorporated. This, because, as mentioned in the theoretical framework, collection and misuse of information are the important factors for privacy concern, whereas improper access and errors of accuracy are outside the scope of the research (see paragraph 2.3.3). This leads to the incorporation of five items, all measured on a 7-point Likert scale anchored with strongly disagree and strongly agree. However, after preliminary interviews, one item was deleted, as it closely resembled another (see paragraph 3.3). The final scale thus only had four items for privacy concerns.

Table 3

Scale for privacy concerns

Item name	Item text	Measurement
PRICON1	I am concerned about my privacy when companies ask me for personal information online	7-point Likert Scale
PRICON2	When companies ask me for personal information online, I sometime think twice before providing it	7-point Likert Scale
PRICON3	I am concerned that companies are collecting too much information about me	7-point Likert Scale
PRICON4	Companies should never share my personal information with other companies unless I authorize it	7-point Likert Scale
PRICON5*		7-point Likert Scale

Note. * PRICON5 has been deleted after preliminary interviews, see 3.3 and table 7.

3.2.4 **IV3: Manipulation Concerns**

As concerns for manipulation is a new empirical construct, in this research, the variable is conceptualized and operationalized for the first time. A new scale was specifically developed for this research, based on *the paradigm for measuring marketing constructs*, outlined by Churchill (1979). Further inspiration is obtained from Mackenzie, Podsakoff & Podsakoff, (2011). A first step in developing new scales is the specification of the construct's domain, or conceptualization in other words (Churchill, 1979).

Based on a literature review, as is defined in the theoretical framework, manipulation concerns of consumers relate to the concerns that online freedom of choice is impeded by advanced and manipulative personalization techniques used by a company to steer the consumer in a desired direction. Therefore, central to manipulation concerns is the perceived freedom of choice a consumer has to make decisions, such as purchasing decisions online. As the degree of online freedom of choice is impeded by steering behaviour from the company, manipulation concerns can also be reflected by the opinion that one is steered towards desired behaviour by a company, such as purchasing a product. As the important second step is the creation of items that clearly capture the specified domain (Churchill, 1979), items developed for the construct therefore clearly have to capture a consumer's perception towards online freedom of choice in an personalization environment, and likewise perception towards steering behaviour by the company is in effect. It should be noted however that items must only entail the perception that a consumer herself can be prone to manipulative efforts. Concerns that other consumers may be influenced, such as elderly people (Newman, 2014) should not be incorporated in the items.

Furthermore, as is reflected in the theoretical framework, manipulative attempts of companies can take several forms. On the one hand, consumers may be shown certain products or services (Calo, 2013; Newman 2014) or different prices (Angwin, 2014). On the other hand, advertisements of products and services shown to all consumers may be differentiated themselves (Kaptein & Eckles, 2010). The scale for manipulation concerns has to take these differences into account. This is done by mentioning personalized offerings instead of separate categories of manipulation, because personalized offerings consist of both personalized products and services as well as personalized prices and personalized advertisements.

Based on the above, a scale has been constructed that consists of four items which are measured using a 7-point Likert scale anchored with strongly disagree and strongly agree. The developed scale can be found in the table below.

Table 4

Scale for manipulation concerns

Item name	Item text	Measurement
MANCON1	I am concerned that my freedom of choice is impeded by personalized offerings	7-point Likert Scale
MANCON2	I think I lose control over my freedom of choice when I receive personalized offerings	7-point Likert Scale
MANCON3	I am concerned that personalized offerings steer me towards behaviour that is beneficial for the company	7-point Likert Scale
MANCON4	I feel that personalized offerings push me towards purchasing certain products and services online	7-point Likert Scale

3.2.5 IV4: Discrimination Concerns

Similarly to manipulation concerns, also the construct for discrimination concerns is a new empirical construct that is conceptualized and operationalized in this research for the first time. Based on Churchill (1979) and Mackenzie et al. (2011) a new scale for this construct is developed in this paper.

From the literature review in the theoretical framework, it becomes clear that discrimination concerns relate to consumer’s concerns of not receiving the same personalized offerings online that other consumers receive. Thus, central to discrimination concerns is the perceived violation of equal treatment online, or in other words: the perceived discrimination compared to other consumers. Important is that consumer’s perceptions are measured, which means that no actual discrimination has to take place. Items should therefore clearly capture the perception of consumers regarding online equal treatment (Churchill, 1979). It has to be remarked that items should only capture a consumer's perception that she herself is prone to online discrimination. Measures regarding individual concerns that other people, or social groups in a broader context are discriminated online are not within the scope of this research.

As is mentioned in the theoretical framework, a specific form of online discrimination that may be of concern to the consumer is price discrimination, which means that consumers are receiving a different price for the same product of service, based on their personal information (Odlyzko, 2003). In addition, it has to be noted that there is a distinction between discrimination that occurs on the basis of a correct customer profile and discrimination that occurs on the basis of an incorrect customer profile (Barocas, 2014). This last form of online discrimination happens for example if a consumer has a behaviour of looking for bargains online and is thus marked as a person with a low income or with a big household and receives offerings on the basis of this profile, whereas this is not the case (Martijn & Tokmetzis, 2016). Finally, relevant to online discrimination concerns is the fact that consumers who are profiled online, may be prone to discrimination in several occasions during their lifetime (Citron & Pasquale, 2014). This element also has to be incorporated in the scale.

This lead to the construction of 6 different items, all measured using a 7-point Likert scale anchored with strongly disagree and strongly agree. After preliminary interviews however, two items were deleted because they did not measure the correct concept (see paragraph 3.3). The final scale can be found below.

Table 5

Scale for discrimination concerns

Item name	Item text	Measurement
DISCON1	It concerns me that I may not receive the same personalized offerings that other consumers receive	7-point Likert Scale
DISCON2	I find it unfair if I receive different personalized offerings compared to other consumers, based on my customer profile	7-point Likert Scale
DISCON3	I am concerned that I may pay a different price for the same product or service compared to other consumers based on my personal information	7-point Likert Scale
DISCON4	I am concerned that once I am assigned a customer profile, I may receive different personalized offerings compared to other consumers in several occasions in the future	7-point Likert Scale

Note. *DISCON5 and DISCON6 have been deleted after preliminary interviews, see 3.3 and table 7.

3.2.6 **Moderator: Perceived Information Control**

From the theory above it becomes clear that perceived information control has been researched as a direct effect on behavioural intentions regarding the use of personalized offerings, or similar constructs. However, research that incorporates perceived information control as a moderating factor governing the relationships between privacy concerns and the likelihood to use personalized offerings, has been much scarcer. Only Taylor et al. (2009) conceptualize information control as a unidimensional construct that covers a general level of perceived information control as a moderating effect. Their scale for the construct involves 4 items measured on a 7-point Likert scale anchored with strongly disagree and strongly agree, which are based on Liu, Marchewka & Ku (2004). However, that scale was used for specific websites, which means that it is less suitable for this research, which touches on perceived information control of information given to companies in general. Instead, this research makes use of a scale of a research where perceived control is a direct effect on the construct of privacy concern, namely the research of Xu, Dinev, Smith & Hart (2008) who, inspired by Xu (2007), constructed 4 items measuring perceived information control on commercial websites (or companies behind them) on a 7-point Likert scale.

Table 6

Scale for perceived information control

Item name	Item text	Measurement
PCONTR1	I believe I have control over who can get access to my personal information collected by companies online	7-point Likert Scale
PCTONR2	I think I have control over what personal information is released by companies	7-point Likert Scale
PCONTR3	I believe I have control over how personal information is used by companies	7-point Likert Scale
PCONTR4	I believe I can control my personal information provided to companies online	7-point Likert Scale

3.3 **Preliminary Interviews**

To assess face validity of the survey, three preliminary interviews with fellow students have been held. These interviews tested whether the constructs are measured correctly by the corresponding items. Assessing face validity was especially important for the newly constructed scales for manipulation concerns and discrimination concerns. The interviews showed that items for these constructs are closely measuring the respective concepts which means that items could be incorporated in the final survey instrument. Overall, with respect to face validity, all constructs proved to be different from each other; that is, interviewees agreed with the scales for each construct and did not find items to belong to different constructs.

However, as mentioned above, based on suggestions from the interviewees, some revisions to the scales have been made. First of all, two items in the scale for discrimination concerns were deleted since they did not measure the same concept as other items did; those items did not focus on the aspect of discrimination but rather on mistakes regarding customer profiling, which is not measured in this research. With regards to other constructs, some additional minor revisions to the items had to be made. The items for value of personalized offerings were revised so that there existed a higher degree of similarity between the items, removing the items that measured the specific value of (time) efficiency of searching online and incorporating additional items from de Pechpeyrou (2009) that measured value of personalized offerings on a more general level. Thus, in the end, the scale for value of personalized offerings is not composed of multiple valid scales (as was planned) but is instead based on one specific scale in the literature, namely the one used by de Pechpeyrou (2009). As a final revision based on the interviews, one item for privacy concerns was deleted from the scale as it too closely resembled another. Remaining items proved to measure the correct construct and further were clear and not too suggestive. Below is a table of revisions to the scales used in this research, based on the preliminary interviews. The original items before revision can also be found in this table.

Table 7

Revisions to scales based on preliminary interviews

Item name	Item before revision	Action	Reason
VALPER2	Personalized offerings increase my searching efficiency	Revised	Scale adapted to be more general
VALPER3	Personalized offerings reduce my time of information searching	Revised	Scale adapted to be more general
PRICON5	When I give personal information to a company for some reason, the company should never use the information for any other reason	Deleted	Item closely resembled PRICON4
DISCON5	I am concerned that companies mistakenly put me in a customer profile	Deleted	Item did not capture concept
DISCON6	I am concerned that if companies mistakenly put me in a customer profile, I may miss out on certain personalized offerings	Deleted	Item did not capture concept

In addition, interviewees assessed what information needed to be given before the questions could be properly answered. There was consensus that only information that delimits terminology used is sufficient, as additional information may steer consumers towards certain answers.

3.4 **Final Survey Instrument**

Based on the preliminary interviews, the final survey instrument has been constructed. An overview of the final survey instrument can be found in appendix II. The structure of the survey is as follows: first, participants are given information about the research topic and specifically about used terminology, namely what personalized offerings and personal information entail. Then, the participant is shown in turn: (1) the two items regarding the likelihood to use personalized offerings, (2) the four items concerning the value of personalized offerings, (3) the four items for privacy concerns, (4) the four items for perceived information control, (5) the four items for manipulation concerns and finally (5) the four items corresponding to discrimination concerns. As has been mentioned before, all items are measured on a 7-point Likert scale, anchored with (1) strongly disagree and (7) strongly agree. Finally, for the purpose of sample description and eliminating possible effects uncontrolled for, the participant is asked some control variables, namely age, gender and highest completed education. An overview of all variables is found below.

Table 8

Overview of variables

Construct	Variable	Measurement	Context
Likelihood to use personalized offerings (proxy: willingness to provide personal information)	LIKPER	2 items on 7-point Likert scale	Combined scale of Chellappa & Sin (2005) and Son & Kim (2008)
Value of personalized offerings	VALPER	4 items on 7-point Likert scale	Based on De Pechpeyrou (2009)
Privacy concern	PRICON	4 items on 7-point Likert scale	Based on Smith et al. (1996)
Manipulation concern	MANCON	4 items on 7-point Likert scale	Newly developed scale
Discrimination concern	DISCON	4 items on 7-point Likert scale	Newly developed scale
Perceived control	PCONTR	4 items on 7-point Likert scale	Based on Xu et al. (2008)
Age	AGE	Ratio between 0-100	Control variable
Gender	MALE	Male=1 Female=0	Control variable
Highest completed education	EDU	1=Lower than High School 2= High School 3= Vocational education (e.g. MBO) 4= Applied education (e.g. HBO) 5= University Bachelor 6=University Master or higher	Control variable

4 Data Analysis and Results

After the final survey instrument has been created, the survey has been distributed among online consumers using the survey software Qualtrics. In this next section, the results of the survey are discussed. As a starting point, the sample used for testing the model and the hypotheses is described. After this follows a description of the preliminary analysis to test the research model used. Finally, and most importantly, the hypotheses are tested based on the collected data. All statistical testing has been performed using IBM SPSS Statistics 25.

4.1 **Sample Description**

The survey software recorded a total of 111 responses. However, some of these responses were only partial and thus non valid. After listwise deletion, which means that the entire record of a respondent was deleted if he had a missing value, the total valid responses of the survey amounted to 108 (*n=108*). With respect to gender, male participation predominated, with 62 respondents (57.41%) being male. Female participation amounted to 46 respondents (42.59%). The mean age of the sample s 25.3 years, with a minimum of 18 and a maximum of 58 (SD= 7,354). A total of 72 respondents (66.87%) have a university degree; 35 respondents (32.41%) completed a university bachelor and 37 respondents completed a university master or higher (43.26%). Furthermore, 23 respondents (21.3%) have completed applied education. Lower education than applied education amounted only to 13 responses (12.04%) with 5 respondents (4.63%) having completed vocational education and 8 respondents (7.41%) having completed high school. None of the respondents had a lower education than high school. An overview of the sample descriptive statistics can be found in the table below:

Table 9

Descriptive Statistics

Construct	Minimum	Maximum	Mean	SD
Likelihood to use personalized offerings	2.00	7.00	4.3739	1.41336
Value of personalized offerings	1.5	7.00	4.0115	1.37873
Privacy concern	1.00	6.00	2.1157	.99850
Manipulation concern	1.00	6.00	3.5625	1.21667
Discrimination concern	1.00	7.00	4.0995	1.37184
Perceived control	1.75	7.00	5.5972	1.10322
Age	18	58	25.30	7.354
Gender	0	1	0.4167	0.49531
Highest completed education	2	6	4.81	1.117

Note. N=108. Gender: 0=Male; 1=Female. Education: 1= lower than High School; 2= High School 3= Vocational Education (e.g. MBO); 4= Applied Education (e.g. HBO); 5=University Bachelor; 6= University Master. All other variables measured on a 7-point Likert scale anchored at 1=strongly disagree and 7=strongly agree.

4.2 **Tests for Reliability and Validity of Constructs**

As a necessary first step in analysis, the reliability and validity of scales have been empirically tested. The preliminary interviews conducted earlier indicated that the scales incorporated in the research appeared to be valid. However, empirical tests were needed to further test the reliability and validity of the used scales, especially of the newly developed scales for manipulation concerns and discrimination concerns.

4.2.1 **Reliability of Scales**

To test for the reliability, or internal consistency of measurements, Cronbach's alpha has been calculated for each scale used in the research. The values can be found in the table below:

Table 10

Cronbach's alpha

Scale	Cronbach's α	Cronbach's α for standardized items
Likelihood to use personalized offerings	0.744	0.744
Value of personalized offerings	0.920	0.920
Privacy concern	0.833	0.832
Manipulation concern	0.817	0.815
Discrimination concern	0.848	0.850
Perceived control	0.867	0.877

As all items have been measured on a 7-point Likert scale, the non-standardized values for Cronbach's alpha can be interpreted, although they do not differ much from the values of alpha for standardized items. The values for Cronbach alpha show that each scale used in the research is sufficiently reliable, as all values exceed 0.7, which is the cut-off point for most research purposes (Allen & Bennett, 2012, p. 217). As the scales for value of personalized offerings, privacy concerns and perceived control are adapted from single researches and then transposed, the alpha coefficient is not surprising. However, the coefficient for likelihood to use personalized offerings shows that also the combined scale of Chellappa & Sin (2005) and Son & Kim (2008) seems reliable. Most importantly however is that the coefficients for manipulation concerns and discrimination concerns show that the newly constructed scales seem reliable as well.

4.2.2 **Construct Validity**

Next, to assess construct validity of all scales, a principal component analysis (PCA) has been performed on all items, using oblique rotation (direct oblimin) at first to assess inter-factor correlation. Inter-factor correlations mostly did not exceed .32, which according to Tabacnick & Fidell (2012) is the cut-off value for using orthogonal rotation, which assumes that factors are independent from each other. After this, another PCA has been performed making use of this orthogonal rotation (varimax). The rotated output of this second PCA can be found below:

Table 11

Rotated component matrix

Item	Component 1	2	3	4	5
LIKPER1	.782				
LIKPER2	.671				
VALPER1	.778				
VALPER2	.894				
VALPER3	.858				
VALPER4	.856				
PRICON1	-.337				.649
PRICON2					.812
PRICON3	-.349				.663
PRICON4					.758
MANCON1				.847	
MANCON2				.847	
MANCON3				.710	
MANCON4				.592	.371
DISCON1			.790		
DISCON2			.848		
DISCON3			.804		
DISCON4			.839		
PCONTR1		.865			
PCONTR2		.895			
PCONTR3		.880			
PCONTR4		.670			

Note. Coefficients lower than absolute value of 0.3 have been suppressed.

As the table shows, five components have been extracted by the PCA, which does not correspond completely with the six different constructs that are used in the research. Unfortunately, the PCA did not distinguish between the dependent variable and value of personalized offerings, but instead grouped them as a single component. The other constructs however, including the newly developed ones, are captured as separate components in the PCA.

The PCA indicated that the constructs for likelihood of using personalized offerings and value of personalization are not completely valid. There may be several reasons for this. First of all, items for the different constructs may have been too similarly written in the questionnaire. However, as there was consensus in the preliminary interviews that the items reflected the correct constructs and these constructs were sufficiently different from each other, this is unlikely. Another reason may be that the construct likelihood to use personalized offerings is composed of too little items for a factor analysis to extract as a unique component. Preferably, a single construct should be composed of at least three items to reflect the complete theoretical domain (Hair et al. 2010, p. 676). In this case, the construct for likelihood to use personalized offerings only consisted of two items. Finally, and most importantly, components may have failed to be extracted because of high correlation between them. In this research, this seems the case, as there is a high correlation between the dependent variable and the value of personalized offerings, namely .718 (see also section 4.3.1 on correlations). This does not necessarily mean that the constructs are not different, the PCA can only not distinguish between them because of the distortion originating from the high correlation.

Besides the reason above, there are other arguments that partially offset the lack of construct validity in the PCA. First of all, the items and scales concerned are not newly developed but are extracted from valid scales in past research and only adjusted a little for the purpose of this research. Therefore, there is some validity to the scales. Secondly, Cronbach's alpha indicated that scales for all constructs proved to be reliable. Thirdly, and already mentioned above, interviewees did distinguish the two constructs and argued that items were substantially different from each other. Finally, the other constructs of the research model, including the newly developed ones, are entirely captured as different components in the PCA and are thus valid.

Because of this, it is decided that the lack of construct validity for the dependent variable and one independent variable is accepted and the research is continued without recollecting the data or altering the research model. This decision is also made for reasons of cost and time efficiency, because it would be too time-consuming to recollect data with the reconstruction of items based on the PCA or the adjusted research model. However, the results from the PCA show a limitation of the research (see also section 5.3, limitations).

4.3 **Hypothesis Testing**

Now that the validity and reliability of the model has been assessed, the next step in the analysis is the testing of the hypotheses. Hypotheses are tested using multiple linear regression. More specifically, a hierarchical multiple regression is conducted following the structure of the conceptual framework mentioned before (section 2.5). The choice for a hierarchical regression instead of a regular multivariate regression is made because the former allows for structurally expanding a previously used model (Field 2013, p. 322). This is the case in this research, as it builds on the existing privacy personalization trade-off and expands it by capturing new concerns and moderation effects, as is shown in the conceptual framework. In addition, hierarchical regression allows for easier comparison between the models, e.g. what happens to the model strength when certain variables are added (Field 2013, p. 324). The hierarchical regression in this research consists of three models: model 1 captures the widely researched effects of the value of personalization and privacy concerns as direct effects (BOX 1), model 2 takes into account the new empirical constructs of manipulation concerns and discrimination concerns (BOX 2) and finally mode 3 includes the moderating effect of perceived information control (BOX 3). The control variables age, gender and education are also included in the regression for the sake of completeness, namely, to include possible effects otherwise unaccounted for. However, if insignificant, they are discarded from the models as they are not part of the explanatory variables in the conceptual framework. This also reduces possible multicollinearity.

4.3.1 Correlations

As a preparation for the linear regression used for testing the hypotheses, first the correlations between the different variables have been assessed. As the regression model (model 3) includes three interaction terms of an independent variable and a moderating variable, it is likely that there are high correlations between the interaction terms and the correspondent independent variables, but also between the interaction terms themselves, as the moderating variable (perceived information control) is the same among all three terms. This may lead to problems of severe multicollinearity and thus results subject to misinterpretation. In fact, when first assessing correlations, interaction terms had many severe correlations between them and with their corresponding independent variables (often $r>.8$, $p<.01$). Therefore, when creating the interaction terms, the variables involved have been centred on the middle value of the Likert scale (4=neither agree nor disagree). This reduces intervariable correlations and hence multicollinearity but does not change the effects. In addition, the coefficients of the model with perceived control as an interaction (model 3) are easier to interpret. This is because after centring, perceived control has a zero value (a Likert scale does not have a zero value), which is important to explain the coefficients of the independent variables (Irwin & McClelland, 2001. See also Field 2013, p. 398-399). Centring on the middle value rather than mean centring has been chosen because centring on the middle value (4 on a Likert scale) gives the easiest interpretation of the coefficient of the independent variable: namely the effect of the independent variable when perceived control is average (neither agree nor disagree), rather than when perceived control is its mean value (which could be well above or below average). See for further clarification paragraph 4.3.3, where the coefficients for each effect are interpreted.

The output of the correlation matrix with centred variables is included in appendix III.

The correlation matrix shows some interesting results. First of all, as mentioned in the section above (4.2.2 on construct validity), the correlation between the dependent variable and value of personalized offerings is highly positive ($r = .718$, $p <.01$), which could explain that the PCA did not distinguish between them when extracting components. Other independent variables seem to significantly correlate with the dependent variable as well, with the exception of discrimination concerns and its interaction term, which have both insignificant correlations. Privacy concerns correlate negatively with the likelihood to use personalized offerings ($r =-.490$, $p <.01$) and so do manipulation concerns ($r=-.314$, $p <.01$). The interaction terms of privacy concerns ($r=-.378$, $p<01$) and manipulation concerns ($r=.314$, $p<0.1$) with perceived information control also correlate negatively with the dependent variable, but to a smaller extent than the main effects, indicating a possible moderation effect. Of the control variables, only age has a significant relationship with the dependent variable. This relationship is positive ($r=.252$, $p<.01$). With respect to multicollinearity, none of the independent variables have correlations with a score above .70. However, even after centring, there appears to be a high correlation between two interaction terms and their corresponding independent variables, namely manipulation concerns and its interaction term ($r=.852$, $p<01$) and discrimination concerns and its interaction term ($r=.891$, $p<01$). This may indicate multicollinearity after all, justifying additional analysis which is discussed in the assumptions section below.

4.3.2 **Assumptions of Linear Regression**

Next, to ensure that the regression analysis is valid, it is imperative that none of the assumptions of linear regression are violated (Allen & Bennett, 2012, p. 182-189).
First of all, it is important that a reasonable ratio between cases to predictors exists. Tabachnick and Fidell (2007) argue that N should ideally be $50+8(k)$. As this research has 5 predictors, N should be 90. Because N is 108 in this research, there exists a reasonable ratio. Second, the 7-point Likert scale variables included in the regression model should approach a normal distribution. If this is the case, they can be included in the regression and the regression becomes valid. The Shapiro-Wilk test for normality indicates that most variables are significantly deviating from normality ($p < 0.05$) (Appendix IV, table 1). According to the

normality test, only the new empirical constructs concerns for manipulation and concerns for discrimination approach a normal distribution, with values of .092 and .255 respectively. However, when assessing normal Q-Q plots and boxplots, it seems that the other variables approach a normal distribution as well (Appendix IV, figures 1-12). The boxplots of privacy concerns and perceived control show a few outliers (figures 6 and 12), but these are not very surprising. This is due to the fact that in this modern era, many consumers are concerned about their informational privacy and similarly perceive low control over personal information. The boxplots show that the answers to the corresponding questions concur with this view (concentrated on agree (=6) for privacy concerns and disagree (=2) for perceived information control). However, there are always, albeit few, individuals who are less concerned about privacy or perceive control over information. But with many answers concentrated on (dis)agree, the differences are larger, hence the outliers.

With respect to the third assumption, the absence of multicollinearity, the previous section on correlations (4.3.1) already mentioned that there were no significant high correlations ($r>0.7$) between the independent variables. However, there were some correlations between interaction terms and corresponding independent variables, even after centring the variables. When including interaction terms in a regression model, some multicollinearity between the interaction term and the independent variables is inevitable. However, severe multicollinearity should be avoided, meaning that tolerance should be >0.1 and VIF values should be <10 (Allen & Bennett, 2012, p. 188). In this research, after centring the variables for interaction terms, this is the case (Appendix IV, table 2).

The final assumption entails normality, linearity and homoscedasticity of residuals. In this research, also this assumption is not violated. When assessing the residual P-P plot, residuals approach a normal distribution (appendix III figure 13). Furthermore, the scatterplot of residuals against predicted values does not indicate non-normality, non-linearity and heteroscedasticity (appendix IV, figure 14).

4.3.3 Linear Regression Output

The hierarchical linear regression consisted of three models with the likelihood to use personalized offerings as the dependent variable. The first model contained the widely researched variables of the personalization privacy trade-off and the control variables. The second model included the new empirical variables. Finally, the third model included the interaction terms based on centred variables. As the control variables were highly insignificant in each of the models, the variables were dropped from the final regression analysis and output. This yielded the following regression model equations:

Model 1: $LIKPER = \beta_0 + \beta_1 VALPER + \beta_2 PRICON$

Model 2: $LIKPER = \beta_0 + \beta_1 VALPER + \beta_2 PRICON + \beta_3 MANCON + \beta_4 DISCON$

Model 3: $LIKPER = \beta_0 + \beta_1 VALPER + \beta_2 PRICON + \beta_3 MANCON + \beta_4 DISCON + \beta_5 PCONTR*PRICON + \beta_6 PCONTR*MANCON + \beta_7 PCONTR*DISCON$

Output of the regression analysis can be found below:

Table 13

Hierarchical regression analysis with LIKPER as the dependent variable

	Model 1			*Model 2*			*Model 3*		
R^2	.545			.553			.583		
Adjusted R^2	.536			.536			.549		
F-value	62.907**			30.192**			17.304**		
Estimates	*β*	*B*	*p*	*β*	*B*	*p*	*β*	*B*	*p*
(Constant)	2.692***		.000	2.756***		.000	2.465***		.009
VALPER	.645***	.626	.000	.636***	.617	.000	.651***	.632	.000
PRICON	-.280**	-.195	.010	-.230*	-.161	.053	-.414***	-.289	.007
MANCON				-.116	-.099	.203	.293	.249	.146
DISCON				.050	.048	.698	-.119	-.114	.450
*PCONTR*PRICON*							.-119*	.232	.062
*PCONTR*MANCON*							.230**	-.433	.022
*PCONTR*DISCON*							.-093	.195	.193

Note. *p<0.1, ** p<0.05, *** p<0.01

The regression output shows that all three models are significant. Model 1, which is the personalization privacy trade-off, has an F value of 62.907 ($p<0.01$) and has an adjusted r-squared of .536, which means that 53.6% of the variation in the dependent variable can be explained by the independent variables. Model 2, which includes the additional concerns of manipulation and discrimination is also significant ($F=30.192$, $p<0.01$) but does not have an increase in adjusted r-squared. The final model, which includes the interaction terms, does have a slightly increased adjusted r-squared, namely .549, and is significant as well ($F=17.304$, $p<0.01$). This indicates that including interaction terms is a good step in enhancing the predictive validity of the model.

The individual effects of the independent variables are assessed based on unstandardized and standardized coefficients, where the unstandardized coefficients indicate the effect on the dependent variable and the standardized coefficient is used to compare effects with each other. Model 1 is used to assess the effects of privacy concerns and value of personalized offerings (BOX1). Model 2 is used to assess the effects of the additional concerns (BOX2). Finally, Model 3, which is the complete model, is used to evaluate the moderating effect of perceived control over information (BOX3).

Below, the validations of the hypotheses are discussed:

BOX1

> **H1:** The likelihood to use personalized offerings is positively influenced by the perceived individual value of these personalized offerings.

Model 1 shows that the effect of value of personalized offerings on the likelihood to use personalized offerings is highly significant ($p<0.01$). The unstandardized beta is .645. This means that for each point of increase in the Likert scale of value of personalized offerings, the increase on the Likert scale for likelihood to use increases with .645. The more a person values personalized offering, the more willing she is to give up information in return, and thus make use of the personalized offering. Therefore, H1 can be accepted.

> **H2:** The likelihood to use personalized offerings is negatively influenced by the individual's concerns for online privacy.

In model 1, also the effect of privacy concerns on the dependent variable is significant ($p<0.05$), which means that there is an effect of privacy concerns on the likelihood to use personalized offerings. The unstandardized coefficient shows that this effect is -.280, meaning that for each point of increase on the Likert scale for privacy concern, the likelihood to use, decreases with .280 point. The more a consumer is concerned about her privacy, the less she is willing to make use of personalized offerings. However, the standardized beta, which compares the strength of direct effects of independent variables on the dependent variable, shows that the effect of privacy concerns on the dependent variable is much smaller (-.195) than the effect of value of personalized offerings (.626), confirming previous research that privacy concerns are largely offset by individual value of personalized offerings (see e.g. Chellappa & Sin, 2005). All in all, H2 can be accepted.

BOX2

> **H3:** The likelihood to use personalized offerings is negatively influenced by the individual's concerns for being manipulated towards certain behaviour online.

Model 2 indicates that although the effect of manipulation concerns on the likelihood to use personalized offerings is negative, with an unstandardized beta of -.119, this effect is insignificant (p=.203). Therefore, although the effect of manipulation concerns is as hypothesized, it cannot be said that a significance relationship exists. As a consequence, H3 is rejected.

> **H4:** The likelihood to use personalized offerings is negatively influenced by the individual's concerns for being discriminated online.

The unstandardized beta of discrimination concerns on the likelihood to use personalized offerings is .050. However, also this effect is insignificant (p=.698) which means that there is no evidence that there exists a relationship between a person's concerns that she is discriminated by personalized offerings and the willingness to make use of these offerings. However, even if there existed a significant relationship, still the effect is not hypothesized, as it is positive instead of negative. Thus, H4 is also rejected.

BOX3

> **H5:** The relationship between concerns for privacy and the likelihood of using personalized offerings is suppressed by the perceived control over information.

From model 3 it follows that the interaction term between privacy concerns and the moderating variable of perceived information control is marginally significant ($p<0.10$), indicating that the moderating variable has an effect on the relationship between privacy concerns and the dependent variable of likelihood to use. The unstandardized coefficient of the interaction is -.119, whereas the coefficient of privacy concerns is -.414.

The interpretation of the coefficients is as follows: in a model including an interaction term, the unstandardized beta of the independent variable is the effect on the dependent variable if the moderating variable is set to zero. After all, the effect of one-point increase in privacy concerns is no longer β_2 but instead $\beta_2 + \beta_5 PCONTR$. Only when perceived control (PCONTR) is zero, the effect on the dependent variable is the unstandardized beta of privacy concerns β_2, in this case -.414. As perceived control is measured on a Likert scale, as the other variables, there is no zero value of the variable. Therefore, and for multicollinearity reasons explained before (section 4.3.1), the variables in the interaction terms have been centred on the middle value (which is 4: neither agree nor disagree). Thus, in this model, one unit increase in the Likert scale for privacy concerns yields to a decrease of .414 point in the Likert scale for likelihood to use personalized offerings, if perceived control is average (neither agree nor disagree). However, if perceived control is high, e.g. strongly agree on the items, which is 3 after centring, the effect of privacy concerns becomes stronger, namely -.771 (-.414 + -.119*3). On the other hand, when perceived control is low, the effect becomes smaller, namely -.057. The effect of perceived information control is visualized in figure 2 below:

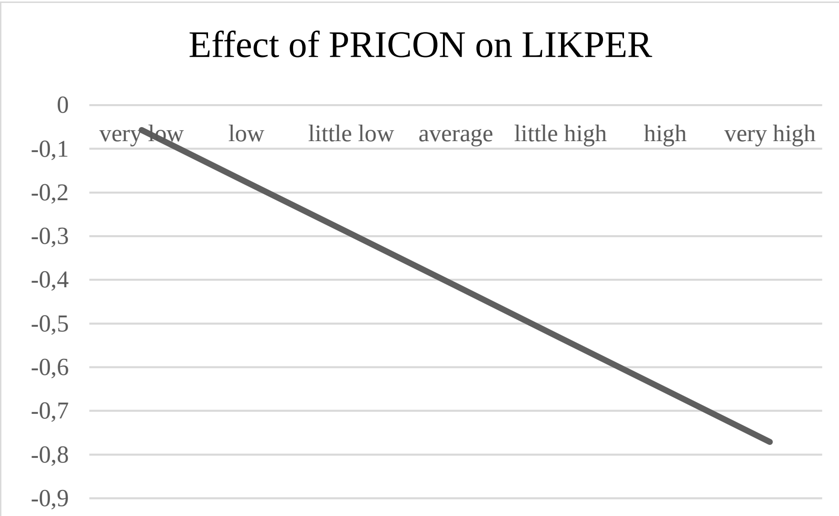

Figure 2. Effect of privacy concerns on the likelihood to use personalized offerings with levels of perceived information control on the x-axis.

The figure indicates that the effect of the perceived control is opposite of what is hypothesized. When perceived information control is high, instead of suppressing the relationship of privacy concerns on the dependent variable, the effect of privacy concerns is stronger. In turn, H5 is rejected.

> **H6:** The relationship between concerns for manipulation and the likelihood of using personalized offerings is suppressed by the perceived control over information.

Model 3 also shows that the interaction term between manipulation concerns and perceived control over information is significant ($p<0.05$). This means that also for manipulation concern, perceived control has a moderating effect. The unstandardized beta of the interaction term is .230, whereas the coefficient of the main effect is insignificant ($p=.146$). Therefore, in this case something interesting happens. When perceived control is average, there is no effect of manipulation concerns on the likelihood to use personalized offerings, since β_3 is insignificant. However, although there is no direct relationship between manipulation concerns and the likelihood to use personalized offerings, there is an effect for certain levels of perceived information control, as β_6 is significant. When perceived control is high or low respectively, there is an effect of manipulation concerns on the dependent variable, and this effect is suppressed by the level of perceived control. For example, when perceived control is moderately high, the effect of manipulation is positive, namely .460 (2*.230). However, when perceived control is very low, the effect of manipulation is -.690 (-3*.230). This effect is visualized below.

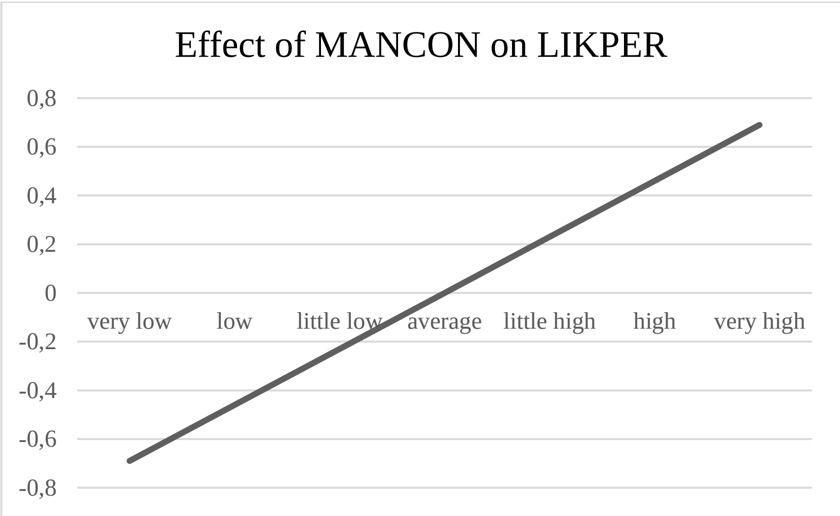

Figure 3. Effect of manipulation concerns on the likelihood to use personalized offerings with levels of perceived information control on the x-axis.

The more control over information a consumer perceives, the less manipulation concerns has a negative effect on the likelihood to use personalized offerings, and the effect can even be offset and become positive by high perceived control. This is a cross-over interaction effect; the effect of manipulation concerns on the likelihood to use personalized offerings is opposite, depending on the level of perceived information control. As a consequence, H6 is accepted.

> **H7:** The relationship between concerns for discrimination and the likelihood of using personalized offerings is suppressed by the perceived control over information.

Finally, as can be seen in model 3, the final interaction term between discrimination concerns and perceived information control is insignificant (p=.193). This means that there is no evidence that perceived control has any effect on the relationship between discrimination concerns and the likelihood to use personalized offerings. As the coefficient of the main effect is also insignificant (p=.450), it cannot even be said that a relationship exists for an average value of perceived information control. As a consequence, the last hypothesis, H7 is rejected.

Below is an overview of the hypotheses accepted and rejected:

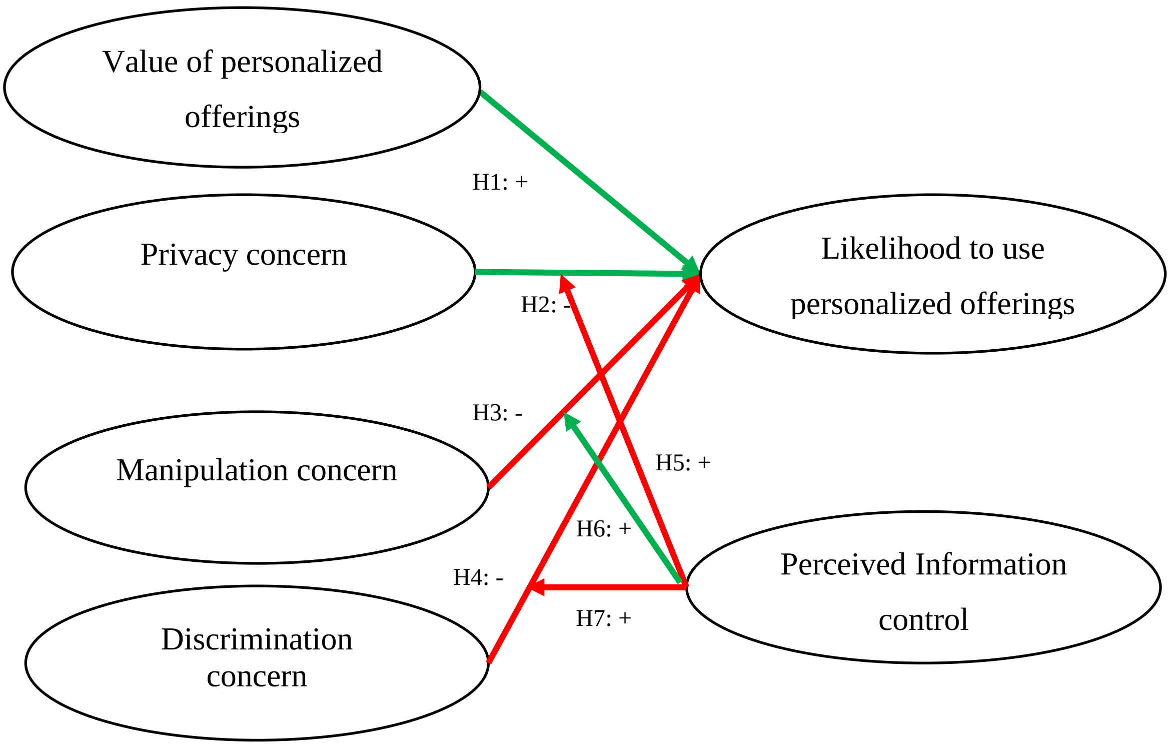

Figure 4. Conceptual framework with hypotheses accepted and rejected.

Table 14

Overview of Hypotheses

Hypothesis	Significance	Effect	Consequence
H1: +	<.01	.645 (+)	Accepted
H2: -	<.05	-.280 (-)	Accepted
H3: -	.203	-.116 (-)	Rejected, insignificant
H4: -	.487	.050 (+)	Rejected, insignificant and wrong effect
H5: +	<.10	-.119 (-)	Rejected, wrong effect
H6: +	<.05	.230 (+)	Accepted
H7: +	.193	-.093 (-)	Rejected, insignificant and wrong effect

5 Conclusion and Discussion

In this conclusion, first the main results of this research are summarized and discussed, after which the managerial and scientific implications of this research are considered. Then, some important limitations are given which may have affected the quality. This conclusion ends with directions for future research on the same or correspondent topics.

5.1 General Discussion

This research tried to update the personalization privacy trade-off that consumers make when faced with online personalization to the modern era of Big Data. In this era, personalization is becoming more sophisticated and consumers do not only have concerns for their privacy but possibly also for companies' attempts of manipulating and discriminating them. Opposed to privacy concern, concerns for manipulation and discrimination have not been extensively researched in academic literature on an empirical level. Rather, literature in fields of marketing, economics, law and psychology have discussed the concerns on a theoretical and more abstract level. Therefore, this research is probably among the first to empirically have tested these concerns. An important aspect of this research was therefore scale construction. Valid scales for manipulation concerns and discrimination concerns have been developed and subsequently applied in a consumer survey. Based on the data from this consumer survey, this research tested several hypotheses related to the updated personalization concerns trade-off. It did so by using a conceptual model that incorporated the personalization privacy trade-off on the likelihood of using personalized offerings and then extended it by capturing the concerns for manipulation and discrimination as part of the negative balance of the trade-off. Finally, this model was amended further by including a moderating effect of perceived information control on the relationship between the concerns and the likelihood to use.

This research confirmed earlier research that there exists a positive relationship between value of personalized offerings and the likelihood to use personalized offerings provided by a company (H1). In addition, this research also confirmed the negative relationship that privacy concerns have on the same likelihood to use personalized offerings (H2), although this effect is much smaller than the positive effect of value of personalized offerings. However, in regard to the possible moderating effect of perceived information control on this relationship, although a significant effect was found, this effect was opposite as was hypothesized (H5). Instead of suppressing the relationship between concerns for privacy and the likelihood to use personalized offerings, perceived information control strengthened the relationship, meaning that consumers with high privacy concern, but who thought were in control over their information, were less likely to accept personalized offerings than consumers with high privacy concerns and low perceived information control. Although this effect seems counterintuitive, there might be some explanation for it. John, Acquisti & Loewenstein (2010) found that privacy policies, for example privacy statements, may actually decrease consumer willingness to provide privacy sensitive information and in turn decrease likelihood to use personalized offerings. This, because the increase in perceived control may backfire by rousing privacy concerns in those instances. For example, a consumer may perceive to have more control over which information she releases to the company, but because of this option, she becomes more concerned that the company uses her information (otherwise the option to control information would not be necessary) and in turn chooses to opt out from personalized offerings. Thus, increases in perceived control may actually lead to a stronger relationship between privacy concerns and the likelihood to make use of personalized offerings.

With respect to the concerns for manipulation, unfortunately, this research found no significant effect on the likelihood to use personalized offerings (H3). Although the effect was as hypothesized, the result was insignificant. However, when the moderating effect of perceived information control is incorporated in the relationship, something interesting happens. For low levels of perceived control over what personal information the company has access to, concerns for being manipulated leads to less likelihood to use personalized offerings. However, for high levels of perceived control, this effect is opposite and manipulation concerns can lead to even more likelihood to use. Thus, there is a high effect of perceived information control on the relationship between manipulation concerns and the dependent variable (H6). This may indicate that consumers who are concerned about being manipulated by companies into buying something only care for this manipulation when they are not in control over the attempt. That is, that they respond negatively towards the offering only when they perceive no control over what personal information this company uses to manipulate their behaviour. With high perceived control, consumers who are concerned they are being steered by companies into buying something, may feel that they have ‘allowed’ the manipulation and end up buying the product more than consumers who perceive information control to be low. Thus, consumers who feel they are in control over the manipulation attempts, may fall for this manipulation easier. This could maybe be explained because it eases decision making; consumers do not have to actively search for products they like because these products are pushed towards them by the company in a personalized fashion. These consumers may see a reduction of their freedom of choice as a bless because they do not have to make effort to choose.

Finally, this research found no effect between discrimination concerns and the likelihood to use personalized offerings (H4) and no moderating effect of perceived information control on this relationship (H7). Thus, it cannot be said that there exists a relationship between concerns for discrimination and the likelihood to use personalized offerings. This may indicate that consumers do not mind being offered different products or a different price for their decision to purchase the different offering. It has to be noted however that the survey only asked consumers for views towards being discriminated themselves and not discrimination in general, for example, concerns that minorities are being discriminated. Were this the case, the result would maybe be different as ethical considerations may surface more predominantly.

All in all, this research confirmed that the personalization privacy trade-off is still valid in the era of Big Data. There exists a tension between the value of personalization and concerns for privacy. The value of personalization has however a larger effect, indicating that a consumer who is concerned about her privacy but also values personalized offerings will still use personalized offerings provided by a company. In addition, this research also found an effect of manipulation concerns on the likelihood of using personalized offerings, but only related to the perceived control over personal information the company has access to. In this sense, also manipulation concerns have a role in the trade-off. When a consumer has concerns for privacy and concerns for manipulation (with low perceived information control), these effects outdo the effect of value of personalized offerings.

5.2 Implication

5.2.1 Scientific Implication

This research provides several implications for academic literature on the topic of online personalization in a commercial environment. First of all, this research has, in line with previous research, confirmed the existing trade-off between value of personalization and concerns for privacy when consumers have to decide to make use of forms of personalization of companies. In this sense, this research has complemented previous research. However, this research also supplemented previous studies as it updated the trade-off to the modern era of Big Data, which entails sophisticated personalization but also more concerns for the consumer than just privacy. This research has tried to take these aspects into account and therefore, the trade-off in this research is more suited for the current time we live in, comparing to older studies on the same subject.

On a deeper level, this research has managed to bridge a gap between theoretical literature on the topic which is common in some research fields, such as the field of law or business ethics, and empirical literature, which is more common in other fields, such as the fields of marketing and psychology. It did so by transposing the two concerns for consumers when faced with modern personalization, namely concerns for manipulation and concerns for discrimination, from a theoretical level to an empirical level, by being among the first to conceptualize them as variables in the personalization concerns trade-off and subsequently use them for hypotheses testing on the basis of statistical analyses. Following the first steps of Churchill (1979) this research successfully created valid scales for the theoretical constructs which can be used in subsequent studies on the topic. The construction of valid scales is one of the most important implications for academic literature this research provides.

Finally, this research is also contributing to academic literature as it is one of the few that incorporates the general concept of information control as a moderating effect in the research model. Previous research mainly assessed the effect of perceived information control as a direct effect on the likelihood to make use of personalization or in a few instances as a moderating effect for a specific type of experiment. Capturing perceived information control in general as a moderating effect on the relationship between concerns and the likelihood to use personalized offerings is possibly unique to this research and it provided for some interesting results. Perceived control can lead to lower likelihood of using personalized offerings when consumers feel manipulated, but it also strengthens the relationship of privacy concerns on the dependent variable whereas it is expected to suppress the relationship. These interesting results are important scientific implications which may require further research.

5.2.2 Practical Implication

As noted above, this research once more confirms the existing trade-off between value of personalization and concerns for privacy and shows that it is relevant more than ever in the age of Big Data. In this age, where commercial companies can use enormous amounts of data of their customers to discover patterns and act on them, personalization becomes very sophisticated, and concerns for privacy live among many consumers. Therefore, managers of companies offering personalized offerings would do good to keep this trade-off in mind. They should not engage in personalization systems to deliver their products to their consumers without keeping in mind the impact on consumer’s privacy. However, this does not mean that managers should abstain implementing personalization systems altogether as this research shows that for a consumer, the value of personalized offerings still prevails in the decision to accept personalized offerings. A consumer who is concerned about her privacy but also values personalized offerings, will likely still gladly make use of the personalized offerings provided by the company, probably because these offerings, originating from highly sophisticated data analysis, provide much value to the consumer in terms of e.g. monetary value or costs of searching. All in all, better still is for managers to maximize the benefits of personalization while at the same time minimizing privacy impact. The upcoming trend, and even obligation (e.g. Article 25 GDPR), of implementing privacy by design provides an effective tool to achieve these goals as privacy is considered from scratch when developing effective personalization systems.

In this sense, managers would do well to broaden their scope to not only take privacy into account in their decision making, but also the effect manipulation may have on their customer's decision making. Managers of commercial companies should minimize exploiting consumer irrationality and attempts to steer consumers towards their products, as this may backfire when consumers perceive that they have low control over the information given to the company, or in more general terms when they are suspicious towards the personalization techniques. On the other hand, high perception of control may lead to more success of these steering attempts as consumers may feel they are in control over the attempts and may benefit from them by having to think less about decision making. Thus, managers should be transparent about the information they collect and give consumers control, or at least a feeling of control (as this research only measured perception), over what information they can release. However, this research also showed that this increase of control may ultimately be maleficent for companies as it may strengthen the effect of privacy concerns on the decision of the consumer to make use of personalized offerings, possibly due to raising suspicion that privacy may be violated. Thus, a clear balance between too little and too much transparency and control has to be made.

Finally, some remarks are to be given to policymakers implementing regulation on this topic. This research shows once more that privacy is of great importance and that privacy protecting regulation needs to be in place in the context of online commercial transactions. However, this regulation should not ultimately hamper the development of sophisticated personalization systems as these are valuable not only to companies but also consumers themselves. Instead, a clear balance between benefits of personalization and privacy has to be found. In addition, regulation should also include protection against unwanted manipulating attempts of companies to steer consumers towards buying products, as these concerns may be important to the consumer when they feel they are not in control over their information. Instead of having several distinct laws for the different concerns, a regulation on fair trade practices may be sufficient as it can designate both privacy violation as well as manipulation as unfair or deceptive trade practices towards consumers. This is already the system in the United States, where no separate privacy regulation, such as the GDPR, exists.

5.3 **Limitations**

Although this research is conducted with careful considerations, there are some limitations to it. Some of which are inherent to the scale of this research and others could have been avoided with hindsight.

A first set of limitations is related to the scale of this research. Due to the fact that this research has been conducted in the setting of a Master thesis, there were some time and cost constraints. This meant inter alia that this research has only a small sample size of *n*=108. Although this sample size was sufficient for a reasonable ratio between cases tot predictors (Tabachnick and Fidell, 2007), the small sample size negatively impacted statistical power of the research which made it less likely to detect significant effects. This is possibly applicable to the effect manipulation concerns have on the dependent variable. Although the direction and size of the effect was as was hypothesized, the small sample size could have rendered the effect insignificant. Due to some insignificant effects, this research has not been able to find conclusive results of some of the relationships proposed.

A second limitation related to the scale of the research is the fact that the survey in this research has only been distributed to consumers living in the Netherlands. Although there are probably similarities to consumers in other countries in Western Europe, in essence the conclusions of this research should be limited to the Netherlands only. Furthermore, another limitation of this research is the fact that it has made use of a survey measuring intent rather than a more experimental setting in which actual behaviour could be measured. Although intent serves as a good proxy for actual behaviour (see e.g. Malhotra et al. 2004), measuring actual behaviour could lead to different results. For example, a consumer may say that he would abstain from personalized offerings because she is concerned about her privacy, but in the end still use the personalized offering in the heat of the moment for example due to spontaneous or impulsive behaviour. These types of behaviours are not captured in intent.

With hindsight however, some considerations, despite the scale of the research, could have been made regarding manipulating certain variables to measure them directly. For example, instead of measuring perceived control, this research could have made use of a more experimental survey in which different situations of information control were simulated, comparing the different groups on the results of the relationship between the concerns and the dependent variable. It has to be noted however that this research intended to measure the variables in general whereas a simulated environment would lead to more specific results related to that environment, for example a simulated website on a specific topic (e.g. travel).

The most important limitation that could have been avoided with hindsight is the fact that the dependent variable likelihood to use personalized offerings and the independent variable value of personalized offerings were perhaps too similar too each other to assess the relationship between them correctly. Although a proxy was used for the dependent variable, namely the willingness to provide personal information, and preliminary interviewees distinguished the variables sufficiently, the PCA still could not extract two separate components from the sets of items, whereas it did so correctly for the other variables. This could have resulted in an overestimated effect, indicated by the high amount of significance of the relationship. Possibly, the phrasing of the items for the different variables were too similar. Another possibility is the fact that the dependent variable was made up of only two items, which is too little for a PCA to extract as a single construct (Hair et al. 2010, p. 676). Therefore, with hindsight, some extra attention had to be given to the conceptualization and item construction for the dependent variable.

5.4 **Directions for Future Research**

Future research is advised to act on the limitations of this research provided above.
First of all, this research could be replicated in the future on a larger scale, with less constraints of time and costs. A larger sample size may improve statistical power and in turn help finding significant effects of the empirical concerns of manipulation and discrimination. With this, some more definite conclusions can possibly be made on the effect that concerns for manipulation and concerns for discrimination actually have on the likelihood to use personalized offerings, and how these concerns are compared to the concerns for privacy and the other side of the trade-off, the value of personalized offerings. Due to the fact that this research has provided new and valid scales for the constructs of these concerns, replication in this sense would not be difficult.

In addition, researchers are implored to replicate the study on a broader scale, asking perceptions of consumers of different countries and comparing results to get an understanding how different consumers react, which could serve as input for local regulating authorities. Furthermore, researchers could conduct similar studies but with a different research design, such as a controlled experiment, to gain understanding of actual behaviour instead of only intent and assess if the results differ significantly or are similar. For example, researchers could measure actual behaviour of providing information to companies or measure information control directly rather than in the form of perception. Also, researchers could amend or build the model used in this research by incorporating other moderating variables related to the personalization privacy trade-off such as the roles of trust and transparency or effects of company reputation. Possibly, also other concerns for the consumer besides privacy, manipulation and discrimination can be distinguished and incorporated in the model or these concerns can be further updated. For example, by assessing general concerns for discrimination rather than individual concerns.

Researchers would also do well to study specific aspects of this research. First of all, differences can be made between the types of personal information, such as contact details or age, financial information or even personal information related to health or race. For each type of information, the trade-off consumers make can be very different. Discrimination concerns may be more prevalent for personal information related to race than for contact details or age. Future research is advised to explore these different trade-offs in more detail. Finally, future research could embark on finding explanations for some interesting results this research provides. For example, answering the question if perceived control really strengthens the relationship between privacy concerns and intent to use personalized offerings and if yes, why that is so. Also, researchers could further explore the relationship that manipulation concerns only affect the intent to use personalized offerings with high or low levels of perceived controls.

All in all, future research on this topic is highly advised as to gain better understanding of the trade-off between personalization and concerns such as privacy, manipulation and discrimination, which is important more than ever in the era of Big Data.

Reference List

Adolphs, C., & Winkelmann, A. (2010). Personalization research in e-commerce-A state of the art review (2000-2008). *Journal of Electronic Commerce Research*, *11*(4), 326-341.

Adomavicius, G., & Tuzhilin, A. (2005). Personalization technologies: a process-oriented perspective. *Communications of the ACM*, *48*(10), 83-90.

Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing*, *91*(1), 34-49.

Alba, J., Lynch, J., Weitz, B., Janiszewski, C., Lutz, R., Sawyer, A., & Wood, S. (1997). Interactive home shopping: consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing*, 38-53.

Alhouti, S., Johnson, C. M., & D'Souza, G. (2016). The Complex Web of Values: The Impact on Online Privacy Concerns and Purchase Behavior. *Journal of Electronic Commerce Research*, *17*(1), 22.

Allen, P. & Bennett, K. (2012). *SPSS Statistics, a practical guide version 20*. Melbourne: Cengage Learning Australia.

Angwin, J. (2014). Dragnet Nation: A Quest for Privacy, Security, and Freedom in a World of Relentless Surveillance: Chapter 1: Hacked. *Colo. Tech. LJ*, *12*, 291-308.

Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer networks*, *54*(15), 2787-2805.

Awad, N. F., & Krishnan, M. S. (2006). The Personalization Privacy Paradox: An Empirical valuation of Information Transparency and the Willingness to be Profiled Online for Personalization. *MIS Quarterly*, 30 (1), 13-28.

Baker, S., Gentry, J. & Rittenburg, T. (2005). Building understanding of the domain of consumer vulnerability. *Journal of Macromarketing*, *25*(2), 128-139.

Barocas, S. (2014). Data mining and the discourse on discrimination. In *Data Ethics Workshop, Conference on Knowledge Discovery and Data Mining*.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *Cal. L. Rev.*, *104*, 671-732.

Bilenko, M., & Richardson, M. (2011). Predictive client-side profiles for personalized advertising. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 413-421.

Bodle, R. (2014). Predictive Algorithms and Personalization Services on Social Network Sites: Implications for Users and Society. In A. Bechman, & S. Lomborg (e.d.), *The Ubiquitous Internet*. London: Routledge.

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2012) Misplaced Confidences: Privacy and the Control Paradox. *Social Psychological and Personality Science 4(3)*, 340-347.

Brill, J. (2013). Reclaim Your Name. *In 23rd Computers Freedom and Privacy Conference.* Washington D.C., June 26, 2013. Retrieved from: https://www.ftc.gov/sites/default/files/documents/public_statements/reclaim-your-name/130626computersfreedom.pdf.

Calo, R. (2011). The boundaries of privacy harm. *Ind. LJ*, *86*, 1131-1162.

Calo, R. (2013). Digital Marketing Manipulation. *George Washington Law Review*, 12, 995-1051.

Campbell, A. J. (1997). Relationship marketing in consumer markets: A comparison of managerial and consumer attitudes about information privacy. *Journal of Interactive marketing*, *11*(3), 44-57.

Chellappa, R.K., & Sin, R.G. (2005). Personalization Versus Privacy: An Empirical Examination Of The Online Consumer's Dilemma Information. *Information Technology and Management,* 6 (3), 181-202.

Churchill Jr, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of marketing research*, 64-73.

Citron, D. K., & Pasquale, F. (2014). The scored society: due process for automated predictions. *Wash. L. Rev.*, *89*, 1-33.

Cukier, K., & Mayer-Schoenberger, V. (2013). The rise of big data: How it's changing the way we think about the world. *Foreign Aff.*, *92*, 28-40.

Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization science*, *10*(1), 104-115.

Culnan, M. J., & Bies, R. J. (2003). Consumer privacy: Balancing economic and justice considerations. *Journal of social issues*, *59*(2), 323-342.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434-449.

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information systems research*, *17*(1), 61-80.

Dupre, E. (2015). Privacy and Security Remain Tall Orders for Today's Marketers. *Direct Marketing News,* (January 23). Retrieved from http://www.dmnews.com/infographics/privacyand-security-remain-tall-orders-for-todays-marketers-infographic/article/393980/]

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897-904.

Field, A. (2013*) Discovering Statistics using IBM SPSS Statistics*. London: Sage Publications.

Fishbein, M., & Ajzen, I. (1977*). Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

FTC (Federal Trade Commission), *Privacy online: Fair information practices in the electronic marketplace*, US Congress, Washington, DC (2000).

FTC (2009). *Self-Regulatory Principles for Online Behavioral Advertising*, FTC staff report, available at: https://www.ftc.gov/reports/federal-trade-commission-staff-report-self-regulatory-principles-online-behavioral (pp. 21-25).

FTC (2012). *Protecting Consumer Privacy in an Era of Rapid Change*, available at: https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf (pp. 18-22).

Garcia-Rivadulla, S., (2016). Personalization vs. privacy: An inevitable trade-off? *International Federation of Library Associations and Institutions*, 42(3), 227-238.

Gartner IT Glossary (n.d.). Big Data. Retrieved from: https://www.gartner.com/it-glossary/big-data.

Granville, K. (2018). Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens. *The New York Times.* Retrieved from: https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html.

Hair, J., Black, W., Babin, B., & Anderson, R. (2010) *Multivariate Data Analysis, 7th edition.* New York: Pearson.

Hajli, N., & Lin, X. (2016). Exploring the security of information sharing on social networking sites: The role of perceived control of information. *Journal of Business Ethics, 133*(1), 111-123

Hannak, A., Soeller, G., Lazer, D., Mislove, A., & Wilson, C. (2014). Measuring price discrimination and steering on e-commerce web sites. In *Proceedings of the 2014 conference on internet measurement conference,* ACM, 305-318.

Helberger, N. (2017). Big Data en het consumentenrecht. In P. Blok (Ed.), *Big Data en het recht.* Den Haag: Sdu Uitgevers.

Hirsch, D., (2015). That's Unfair! Or is it? Big Data, Discrimination and the FTC's Unfairness Authority. *Ky. LJ* 103 (2014), 345-361.

Ho, S. Y., & Kwok, S. H. (2002). The attraction of personalized service for users in mobile commerce: an empirical study. *ACM SIGecom Exchanges*, *3*(4), 10-18.

Ho, S., & Kwok, S. (2003). The attraction of personalized service for users in mobile commerce. *ACM SIGecom Exchanges,* 3(4), 10-18.

Ho, S., & Tam, K. (2005). An empirical examination of the effects of Web personalization at different stages of decision making. *International Journal of Human- Computer Interaction,* 19(1), 95-112.

Holbrook, M. B., & Batra, R. (1987). Assessing the role of emotions as mediators of consumer responses to advertising. *Journal of consumer research*, *14*(3), 404-420.

Howard, D., and Kerin, Roger A. (2004) The effects of personalized product recommendations on advertisement response rates: the ‘Try This. It Works!’ technique. *Journal of Consumer Psychology*, 14, 271–279.

Hutchinson, A. (2017). On Facebook’s emotional Ad targeting, the Manipulation of Younger Users, and the concerns of Big Data [Blog post]. Retrieved from: https://www.socialmediatoday.com/social-networks/facebooks-emotional-ad-targeting-manipulation-younger-users-and-concerns-big-data

Jai, T. M. C., & King, N. J. (2016). Privacy versus reward: Do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers?. *Journal of Retailing and Consumer Services*, *28*, 296-303.

John, L. K., Acquisti, A., & Loewenstein, G. (2010). Strangers on a plane: Context-dependent willingness to divulge sensitive information. *Journal of consumer research*, *37*(5), 858-873.

Kaptein, M., & Eckles, D. (2010). Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling. *International Conference on Persuasive* Technology, 82-93.

Kobsa, A. (2007). Privacy-enhanced web personalization. In *The adaptive web*, Springer-Verlag, 628-670.

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online social networks: Why we disclose. *Journal of Information Technology*, *25*(2), 109-125.

Laufer, R. S., & Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of social Issues*, *33*(3), 22-42.

Lee, C. H., & Cranage, D. A. (2011). Personalisation–privacy paradox: The effects of personalisation and privacy assurance on customer responses to travel Web sites. *Tourism Management*, *32*(5), 987-994.

Li, T., & Unger, T. (2012). Willing to pay for quality personalization? Trade-off between quality and privacy. *European Journal of Information Systems*, *21*(6), 621-642.

Li, Y. (2011). Empirical studies on online information privacy concerns: Literature review and an integrative framework. *CAIS*, *28*, 28, 453-497.

Liu, C., Marchewka, J. T., & Ku, C. (2004). American and Taiwanese perceptions concerning privacy, trust, and behavioral intentions in electronic commerce. *Journal of Global Information Management (JGIM)*, *12*(1), 18-40.

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, *35*(2), 293-334.

Mai, J. E. (2016). Big data privacy: The datafication of personal information. *The Information Society*, *32*(3), 192-199.

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*, *15*(4), 336-355.

Martijn, M., & Tokmetzis, D. (2016). *Je hebt wél iets te verbergen.* Amsterdam: De correspondent.

Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. *Journal of Marketing, 81*(1), 36-58.

Mattioli, D. (2012, August 23). On Orbitz, Mac Usres Steered to Pricier Hotels. *The Wallstreet Journal.* Retrieved from: https://www.wsj.com/articles/SB10001424052702304458604577488822667325882.

Meinert, D. B., Peterson, D. K., Criswell, J. R., & Crossland, M. D. (2006). Privacy policy statements and consumer willingness to provide personal information. *Journal of Electronic Commerce in Organizations*, *4*(1), 1-17.

Mothersbaugh, D. L., Foxx, W. K., Beatty, S. E., & Wang, S. (2012). Disclosure antecedents in an online service context: The role of sensitivity of information. *Journal of Service Research*, *15*(1), 76-98.

Newman, N. (2014). How Big Data Enables Economic Harm to Consumers, Especially to Low-Income and Other Vulnerable Sectors of the Population. *Journal of Internet,* 18(23).

Nowak, G. J., & Phelps, J. (1995). Direct marketing and the use of individual-level consumer information: Determining how and when “privacy” matters. *Journal of Direct Marketing*, *9*(3), 46-60.

Odlyzko, A. (2003). Privacy, economics, and price discrimination on the Internet. In *Proceedings of the 5th international conference on Electronic commerce*, ACM, 355-366.

Pavlou, P.A., & Stewart, D.W. (2000) Measuring the effects and effectiveness of interactive advertising: a research agenda. *Journal of Interactive Advertising*, 1, 1–27.

de Pechpeyrou, P. (2009). How consumers value online personalization: a longitudinal experiment. *Direct Marketing: An International Journal*, *3*(1), 35-51.

Peppers, D., Rogers, M., & Dorf, B. (1999). Is your company ready for one-to-one marketing. *Harvard business review*, *77*(1), 151-160.

Phelps, J., Nowak, G., & Ferrell, E. (2000). Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy & Marketing*, *19*(1), 27-41.

De Rek, W. (2017, December 30). Als de big data heersen kan de mens bij het vuilnis. *De Volkskrant*. Retrieved from: https://www.volkskrant.nl/boeken/als-de-big-data-heersen-kan-de-mens-bij-het-vuilnis~a4551594/.

Richards, N. M., & King, J. H. (2014). Big data ethics. *Wake Forest L. Rev.*, *49*, 393-432.

Russom, P. (2011). Big data analytics. *TDWI best practices report, 4th quarter*, *19*(4), 1-34.

Schrage, M. (2014). Big Data's dangerous new era of discrimination. *Harvard Business Review*.

Schumann, J. H., von Wangenheim, F., & Groene, N. (2014). Targeted online advertising: Using reciprocity appeals to increase acceptance among users of free web services. *Journal of Marketing*, *78*(1), 59-75.

Shaw, M. J., Subramaniam, C., Tan, G. W., & Welge, M.E. (2001). Knowledge management and data mining for marketing. *Decision Support Systems*, 31, 127-137.

Sheehan, K. B., & Hoy, M. G. (2000). Dimensions of privacy concern among online consumers. *Journal of public policy & marketing*, *19*(1), 62-73.

Sheng, H., Nah, F. F. H., & Siau, K. (2008). An experimental study on ubiquitous commerce adoption: Impact of personalization and privacy concerns. *Journal of the Association for Information Systems*, *9*(6), 344.

Shiller, B.R. (2014). First degree price discrimination using big data. *Working papers Brandeis University*, *58*, 1-35.

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: measuring individuals' concerns about organizational practices. *MIS quarterly*, 167-196.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: an interdisciplinary review. *MIS quarterly*, *35*(4), 989-1016.

Son, J. Y., & Kim, S. S. (2008). Internet users' information privacy-protective responses: a taxonomy and a nomological model. *MIS quarterly*, 503-529.

Stewart, K. A., & Segars, A. H. (2002). An empirical examination of the concern for information privacy instrument. *Information systems research*, *13*(1), 36-49.

Sutanto, J., Palme, E., Tan, C. H., & Phang, C. W. (2013). Addressing the Personalization-Privacy Paradox: An Empirical Assessment from a Field Experiment on Smartphone Users. *Mis Quarterly*, *37*(4).

Sweeney, L. (2013). Discrimination in online ad delivery. *Queue*, *11*(3), 10-19.

Tabacnick, B. & Fidell, S. (2012). *Using Multivariate Statistics.* Boston, MA: Pearson Education.

Taylor, D.G., Davis, D.F., & Jillapalli, R. (2009). Privacy Concern and Online Personalization: The Moderating Effects of Information Control and Compensation. *Electronic Commerce Research*, 9 (3), 203-223.

Tene, O., & Polonetsky, J. (2012). Big data for all: Privacy and user control in the age of analytics. *Nw. J. Tech. & Intell. Prop.*, *11*, 240-273.

TRUSTe (2016). 2016 TRUSTe/NCSA Consumer Privacy Infographic – US Edition. *TRUSTe*, (December 17-22). Retrieved from: https://www.trustarc.com/resources/privacy-research/ncsa-consumer-privacy-index-us/.

Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research*, 51(5), 546-562.

Wedel, M., & Kannan, P.K. (2016). Marketing Analytics for Data-Rich Environments. *Journal of Marketing*, 80, 97-121.

West, P. M., Ariely, D., Bellman, S., Bradlow, E., Huber, J., Johnson, E., ... & Schkade, D. (1999). Agents to the Rescue?. *Marketing letters*, *10*(3), 285-300.

Westin, A.F. (1967). *Privacy and Freedom.* New York, NY: Atheneum,.

White, T. B. (2004). Consumer disclosure and disclosure avoidance: A motivational framework. *Journal of Consumer Psychology*, *14*(1-2), 41-51.

World Economic Forum Report (2013). *Unlocking the Value of personal data: From Collection to Usage*. Retrieved from: http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf

Xu, H. (2007). The effects of self-construal and perceived control on privacy concerns. *ICIS 2007 proceedings*, 125.

Xu, H., Dinev, T., Smith, H. J., & Hart, P. (2008). Examining the formation of individual's privacy concerns: Toward an integrative view. *ICIS 2008 proceedings*, 6.

Xu, H., Luo, X. R., Carroll, J. M., & Rosson, M. B. (2011). The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision support systems*, *51*(1), 42-52.

Yu, J., & Cude, B. (2009). 'Hello, Mrs. Sarah Jones! We recommend this product!' Consumers' perceptions about personalized advertising: comparisons across advertisements delivered via three different types of media. *International Journal of Consumer Studies*, 33, 503-514.

Appendix I: Literature Review

Study	Areas of Focus	Research design	Key findings/aspects
Box1: Personalization Privacy trade-off			
Smith, Milberg & Burke (1996)	Dimensions of privacy concerns. Scale construction.	Theoretical, interviews, survey (validation of scale)	Privacy concerns consist of the following four dimensions: 1) collection, 2) unauthorized secondary use, 3) improper access and 4) Data Errors.
Culnan and Armstrong (1999)	Information privacy, Trust, Willingness to expose information, Procedural fairness.	Survey	Decision to expose personal information dependent on privacy calculus.
Malhotra, Kim & Agarwal (2004)	Information privacy concerns. Scale construction. Causal model. Willingness to provide personal information.	Theoretical. Survey experiments.	Privacy concern is a second-order construct consisting of collection, control and awareness. Collection is only fair to consumers when they are granted control and awareness over information.
Chellapa and Sin (2005)	Personalization, Privacy, Trust, Online transactions, Likelihood of using personalized services (Willingness to give information as a proxy).	Survey	Relationship between value and concerns (correlation with trust building factors) on likelihood to use personalized services. Personalization value outweighs privacy concerns.
Dinev and Hart (2006)	Privacy Calculus, Trust, Internet Risk, Personal interest, Willingness to provide information.	Survey	Willingness to provide information dependent on privacy concern, trust and personal internet interest.
Awad and Krishnan (2006)	Personalization-privacy paradox, Willingness to be profiled online (service, advertising), Information transparency, Privacy concerns.	Survey	Coining of a paradox. More transparency leads to lower consumer willingness to be profiled. Consumers are more willing to be profiled for services rather than for advertisements.
Kobsa (2007)	Personalization and privacy concerns. Overview of Constructs. Recommendations.	Theoretical	Tension between personalization and privacy concerns. There are several approaches that align personalization with privacy.
Sheng, Nah & Siau (2008)	Personalization, Privacy Concern, intention to adopt of U-commerce, context.	Experiment: scenarios (personalization and context)	Relationship between personalization and privacy concerns in a context of U-commerce. Intention to adopt situational dependent.
De Pechpeyrou (2009)	Attitude towards online personalization. Relationship between personalization and consumer response. Liabilities of personalization. Recommendation.	Longitudinal experiment	Personalization is positively related with response and negatively related with liabilities. Recommendation reminds consumers of product.

Yu and Cude (2009)	Personalized advertising, General Perceptions, Privacy, Advertising effect, Consumer behaviour. Offline mail, Telephone calls, Online mail.	Survey	Negative general perception of personalized advertising. Often immediately rejected. Negative relationship between privacy and personal advertising. Low purchase intentions.
John, Acquisti & Loewenstein (2010)	Context of environment, Privacy concern, Willingness to divulge information.	Field experiments	Contextual information leads to difference in privacy valuations and disclosure. Priming with a privacy statement decreases disclosure.
Hun Lee and Cranage (2010)	Personalization, Privacy assurance, Privacy Concern, Willingness to Give up information, Willingness to adopt services.	Experiment: Travel websites	Personalization perceived as useful. Privacy concerns not necessarily negatively associated with personalization.
Xu, Luo, Carroll, Rosson (2011)	Personalization-privacy paradox, location aware marketing, covert personalization, overt personalization.	Experiment: scenarios (covert and overt personalization)	Personalization privacy trade-off in location aware marketing. Influence of personalization depends on type of personalization (covert and overt) as well as personal characteristics.
Mothersbaugh, Foxx, Beatty & Wang (2012)	Online disclosure, privacy concern, sensitivity of information, website customization benefits.	Experiment: Website scenarios (personalization benefits, information control)	The more sensitive personal information is asked, the less likely a consumer is willing to expose personal information. Perceived benefits can overcome this when concerns are lower or control is higher.
Li & Unger (2012)	Personalization, privacy, quality, information disclosure, willingness to use personalization.	Experiment: website scenarios (privacy, quality of personalization, industry domain)	Relationship between personalization quality and privacy concerns. The better quality personalization has, the more likely a consumer is to use personalization.
Schumann, Wagenheim and Groene (2014)	Social norms, Reciprocity, Privacy concern, Advertising Effectiveness.	Field Study, Online Experiments, Survey	Customers accept targeted advertising in exchange for free services. Advertising form of online currency to repay for benefits.
Jai & King (2016)	Consumer loyalty, Consumer data, information privacy, data brokers, willingness to provide personal information.	Survey	Willingness to provide personal information to data brokers and third party advertisers. Commitment and loyalty lead to more willingness.
Wedel and Kannan (2016)	Big Data personalization, Marketing Mix, Privacy. Predictive personalization.	Theory	New forms of marketing analytics with Big Data. New methods to apply> new forms of personalization.
Garcia-Rivadulla (2016)	Online privacy, personal data, personalization, Privacy Paradox.	Theory	Updated privacy concerns. Combination of personalization and privacy should be possible. Transparency is important.
BOX2: New empirical constructs			
Odlyzko (2003)	Privacy, commercial pricing, price discrimination on the internet, ways of price discrimination.	Theoretical	Exposition of price discrimination on the internet based on personal information. Ways of price discrimination. Incentives

			and negative consequences for the consumer.
Kaptein & Eckles (2010)	Persuasion Profiling, Concerns adaptive persuasion, Ethics online profiling.	Theoretical	Description of concerns related to individual influence strategies. Online manipulation.
Calo (2013)	Digital Market Manipulation, Online manipulation, Online privacy, Vulnerability.	Theoretical	Manipulation of Markets in the digital era. Vulnerability for persuasion and Privacy harms as negative result for the consumer.
Sweeney (2013)	Online advertising, racial discrimination, data privacy, information retrieval, search engine marketing.	Theoretical. Field study: Analyzing Google SEA UI.	On Google search engine advertisements, there exists racial discrimination that shows people certain suggestive advertisements (including the word arrest e.g.).
Angwin (2014)	Privacy and security, Big Data, bias, manipulation.	Theoretical	Overview of negative consequences of Big Data for the consumer, including (financial) manipulation and price discrimination.
Newman (2014)	Economic Harm of Big Data, Vulnerable sectors of population.	Theoretical	Big Data personalization has several negative consequences: racial profiling, vulnerability of consumers groups and price discrimination.
Barocas (2014)	Data mining, discrimination, online equality, fairness, algorithms.	Theoretical	Exposition of discrimination due to data mining. Types of discrimination, principles that are contravened, mechanisms of discrimination.
Citron & Pasquale (2014)	Big Data, Predictive algorithms, automated predictions, stigmatization.	Theoretical	Big Data can lead to automatic predictions i.e. scoring systems in many aspects of society. Regulation on fairness and accuracy is needed for protection of consumers.
Hannak, Soeller, Lazer, Mislove, Wilson (2014)	Personalization, e-commerce, price discrimination.	Field study: Analyzing e-commerce sites (custom scripts)	Assessing degree of price discrimination on big e-commerce websites. 9 out of 10 make use of some form of price discrimination.
Hirsch (2015)	Big Data Discrimination, (Un)fairness, Benefits and Threats, FTC.	Theoretical	Big Data discrimination. Examples given. Responsible big data practice. Unfairness from a FTC viewpoint.
Barocas & Selbst (2016)	Data mining, discrimination, unintentional, intentional, regulatory aspects.	Theoretical	There is a difference between unintentional and intentional discrimination on the basis of data mining, but consequences are the same for consumers.
BOX3: Information control			
Nowak & Phelps (1995)	Personalization, privacy concern, consumer information, alleviating tactics, opt-out.	Theoretical	Theoretical overview of privacy concerns and personalization. Domains of privacy concerns. Alleviating tactics. Opt-out helps alleviating privacy concerns.

Meinert, Peterson, Crisswell & Crossland (2006)	e-commerce privacy, trust, privacy policy statements, web site content, willingness to provide information.	Survey	Information protected as stated in privacy statements has an influence on the willingness to provide personal information. The more privacy is guaranteed, the more likely a consumer is willing to give information.
Xu, Dinev, Smith & Hart (2008)	Privacy concerns, privacy assurance, privacy risk, privacy control. Integrated view of privacy concerns.	Survey	Privacy constructs relate to each other. Instruction, risk and perceived control affect privacy concerns. Control and risk are important factors.
Taylor, Davis & Jillapalli (2009)	Trust, Privacy Concern, Compensation, Information Control, Behavioural intentions (act on desired behaviour).	Experiment: website scenarios (type of information disclosure, cash compensation)	Trust reduces privacy concerns. Privacy concerns on behavioural intentions is moderated by perceived information control, not necessarily by compensation.
Krasnova, Spiekermann, Koroleva & Hildebrand (2010)	Privacy, information disclosure, motivation, online social networks.	Survey	Effect of perceived control over information and trust on perceived privacy risk and self-disclosure of information in turn. Perceived control mitigates relationship between privacy risk and willingness to self-disclose.
Sutanto et al. (2013)	Personalization-Privacy paradox, Mobile advertising applications, Use and gratification.	Field study: mobile apps	Privacy-safe solution for delivering personalized advertising increases use of application and saving of advertisements.
Tucker (2014)	Customer privacy controls, Targeted ads, Personalized ads.	Field study: Facebook	For a non-profit, people responded more favourably to personalized ads when they had the ability to control their privacy settings.
Hajli & Lin (2016)	Privacy Risk, Information sharing, perceived control, social networking sites.	Survey	Influence of users' perceived control of information over information-sharing behaviours on social networks. Perceived control is negatively related to perceived privacy risk and influences information sharing behaviour.
Martin, Borah and Palmatier (2017)	Data breach, Consumer Vulnerability, Privacy, Big Data. Customer Behavior (switching, negative word of mouth). Transparency and Control.	Experiment: scenarios (access vulnerability, transparency, control)	Customers respond negatively to firms collection and use of data. Control and transparency mitigate damaging effect of data vulnerability (breach but also privacy concern) on consumer behaviour (and firm performance as a result).

Appendix II: Final Survey Instrument

Dear participant,

Thank you for participating in this survey, which is used to obtain my Master's degree in Marketing at the Erasmus University in Rotterdam.

The purpose of this survey is to understand consumers' perceptions about online personalization and the benefits and risks involved. Participation in this study will only take around 7 minutes of your time!
I would like to clarify that participation is completely anonymous and your data will be analysed and applied confidentially. Furthermore, there are no right or wrong answers. Therefore, I would like to ask you to fill in the questions truthfully.

Should you have any questions or concerns, please do not hesitate to contact me on my emailadress: 483226dp@student.eur.nl

Thank you in advance for participating, your help is highly appreciated!

Daan van Pinxteren

***

Before you begin, please read the following text carefully:

Commercial companies use more and more information from consumers to give them personalized offerings online. This often happens by assigning customer profiles to consumers with certain personal traits. Currently, customer profiles can be very specific, with some companies having thousands of different profiles. Each of these profiles receives different personalized offerings.

Personalized offerings are those offerings that involve personalization with regards to the purchase of products and services. This may be personalization of (aspects of) products and services themselves, such as a different product design or service package for different consumer groups, but also personalization of the purchasing process, such as personalized systems that recommends certain products and services, personalized discounts, or personalized advertisements for the same products and services, where each customer profile receives a different advertisement with for example a different advertisement headline.

When you read the term personalized offerings in this survey, please keep in mind that this may entail all of the above.

The information that commercial companies collect in order to create customer profiles and in turn to give personalized offerings can be split into the following three categories:

1) Personally identifiable information such as name, address and bank account information

2) Personally unidentifiable information such as age, occupation and gender

3) Anonymous information such as IP-address

The second and third categories are not personal information by itself, but when analysed together can identify individuals. Therefore, these categories can also be seen as personal information.

When you read the term personal information in this survey, please keep in mind that this may entail information from all of the described categories.

The following questions ask you about your perception towards online personalized offerings based on personal information

- *All items measured on a 7-point Likert scale, anchored with 1= strongly disagree 2= disagree 3= somewhat disagree 4= neither agree nor disagree 5= somewhat agree 6= agree 7= strongly agree*

***

DV: Willingness to provide personal information in return for personalized offerings (proxy for likelihood to use)

LIKPER1: I am comfortable providing personal information to companies in return for personalized offerings.

LIKPER2: I refuse to receive personalized offerings when I have to give personal information in return

***

IV1: value of personalized offerings (or attitude towards personalized offerings)

VALPER1: I think that personalized offerings suit my needs

VALPER2: I like offerings from companies that are personalized for me

VALPER3: I have positive feelings towards personalized offerings

VALPER4: I value offerings from companies that are personalized for my preferences

***

IV2: privacy concern

PRICON1: I am concerned about my privacy when companies ask me for personal information

PRICON2: When companies ask me for personal information, I sometimes think twice before providing it.

PRICON3: I am concerned that companies are collecting too much information about me

PRICON4: Companies should never share my personal information with other companies unless I authorize it

***

MOD: perceived information control

PCONTR1: I believe I have control over who can get access to my personal information collected by companies online

PCONTR2: I think I have control over what personal information is released by companies

PCONTR3: I believe I have control over how personal information is used by companies

PCONTR4: I believe I can control my personal information provided to companies online

***

IV3: manipulation concern

MANCON1: I am concerned that my freedom of choice is impeded by personalized offerings

MANCON2: I think I lose control over my freedom of choice when I receive personalized offerings

MANCON3: I am concerned that personalized offerings steer me towards behaviour that is beneficial for the company

MANCON4: I feel that personalized offerings push me towards purchasing certain products and services online

***

IV4: discrimination concern

DISCON1: It concerns me that I may not receive the same personalized offerings that other consumers receive

DISCON2: I find it unfair if I receive different offerings compared to other consumers, based on my customer profile

DISCON3: I am concerned that I may pay a different price for the same product or service compared to other consumers based on my personal information.

DISCON4: I am concerned that once I am assigned a personal profile, I may receive different personalized offerings compared to other consumers in several occasions in the future

***

Control variables

AGE: what is your age? Ratio (Text block)

Gender: what is your gender? Male or Female

Education: what is your highest completed education? 1=Lower than High School 2= High School 3= Vocational education (e.g. MBO) 4= Applied education (e.g. HBO) 5= University Bachelor 6=University Master or higher

Appendix III: Correlation Matrix

Table 12

Pearson correlation matrix

Variable	*LIKPER*	*VALPER*	*PRICON*	*MANCON*	*DISCON*	*PCONTR PRICON*	*PCONTRM ANCON*	*PCONTRDI SCON*	*GEN*	*AGE*	*EDU*
LIKPER	(-)	.718***	-.490***	-.323***	-.009	-.378***	-.314***	.084	.111	.252***	.042
VALPER	.718***	(-)	-.471***	-.263***	-.008	-.436***	-.236**	.070	.139	.288***	-.088
PRICON	-.490***	-.471***	(-)	.470***	.151	.591***	.350***	.058	-.061	-.152	-.085
MANCON	-.323***	-.263***	.470***	(-)	.284***	.378***	.852***	.223**	-.199**	0.13	-.217**
DISCON	.-009	-.008	.151	.284***	(-)	.020	.221**	.891***	-.216**	-0.95	-.067
PCONTR*PRICON	-.378***	-.436***	.591***	.378***	.020	(-)	.599***	.019	-.103	-.141	-.019
PCONTR*MANCON	-.314***	-.236**	.350***	.852***	.221**	.599***	(-)	.223**	-.203**	-.017	-.135
PCONTR*DISCON	.084	.070	.058	.223**	.891***	.019	.223**	(-)	-.225**	-.065	-.056
GENDER	.111	.139	-.061	-.199**	-.216**	-.103	-.203**	-.225**	(-)	-.006	.150
AGE	.252***	.288***	-.152	0.13	-.095	-.141	-.017	-.065	-.006	(-)	-.062
EDUCATION	.042	-.088	-.085	-.217**	-.067	-.019	-.135	-.056	.150	-.062	(-)

Note. ** p <.05 *** p <.01 (two-tailed).

Appendix IV: Assumptions

Table 13

Shapiro-Wilk test for normality

Variable	Statistic	df	Sig.
LIKPER	.951	108	.001
VALPER	.950	108	.000
PRICON	.856	108	.000
MANCON	.979	108	.092
DISCON	.985	108	.255
PCONTR_CEN	.914	108	.000

Table 14

Scores for collinearity

Model	Variable	Tolerance	VIF
MODEL 1	VALPER	.778	1.286
	PRICON	.778	1.286
MODEL 2	VALPER	.770	1.300
	PRICON	.647	1.545
	MANCON	.729	1.372
	DISCON	.912	1.097
MODEL 3	VALPER	.716	1.398
	PRICON	.380	2.633
	MANCON	.145	6.899
	DISCON	.185	5.400
	PCONTR*PRICON	.276	3.617
	PCONTR*MANCON	.121	8.263
	PCONTR*DISCON	.189	5.283

Note. Dependent variable: LIKPER

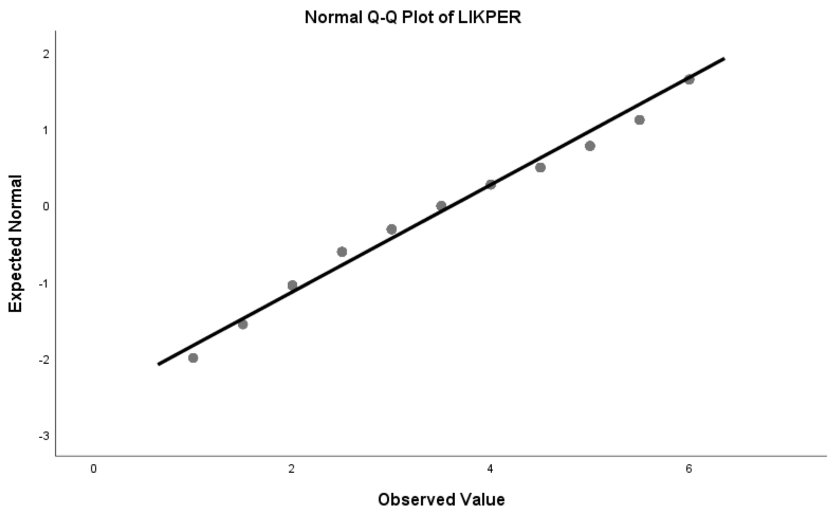

Figure 5. Normal Q-Q plot of LIKPER

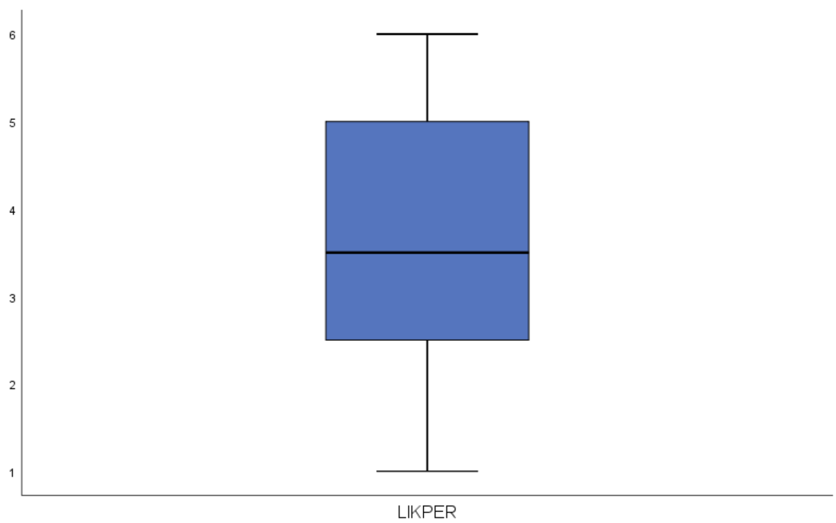

Figure 6. Boxplot of LIKPER

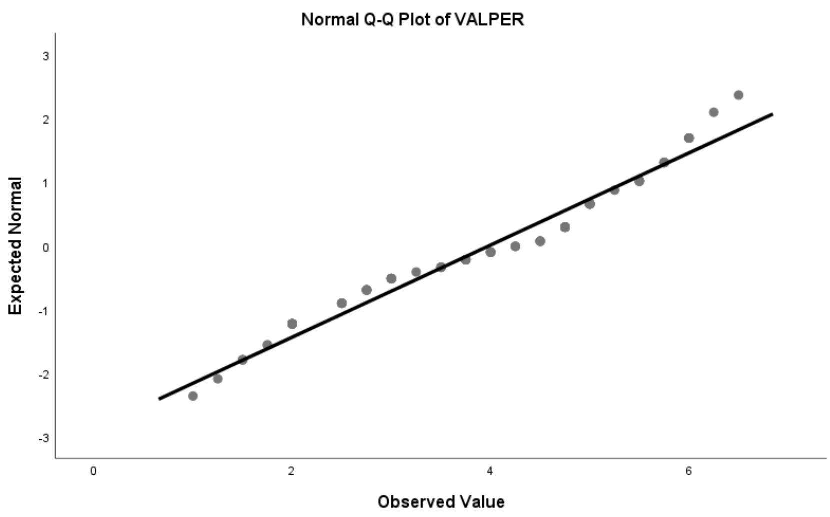

Figure 7. Normal Q-Q plot of VALPER

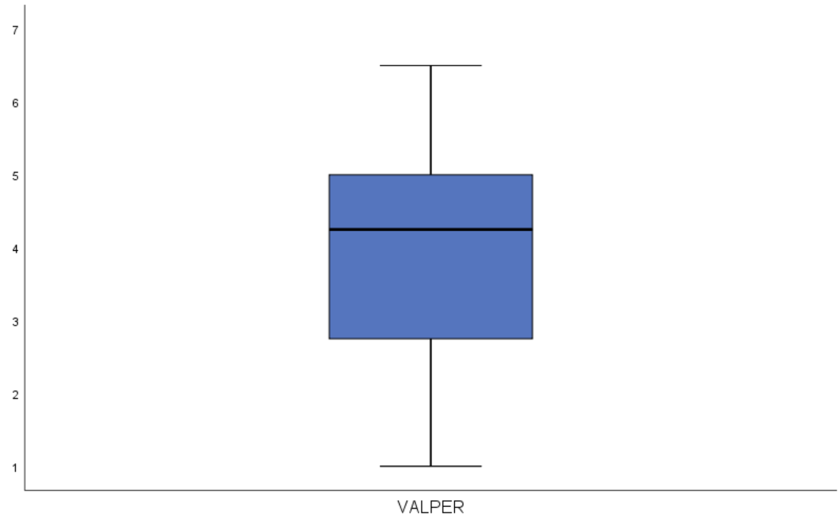

Figure 8. Boxplot of VALPER

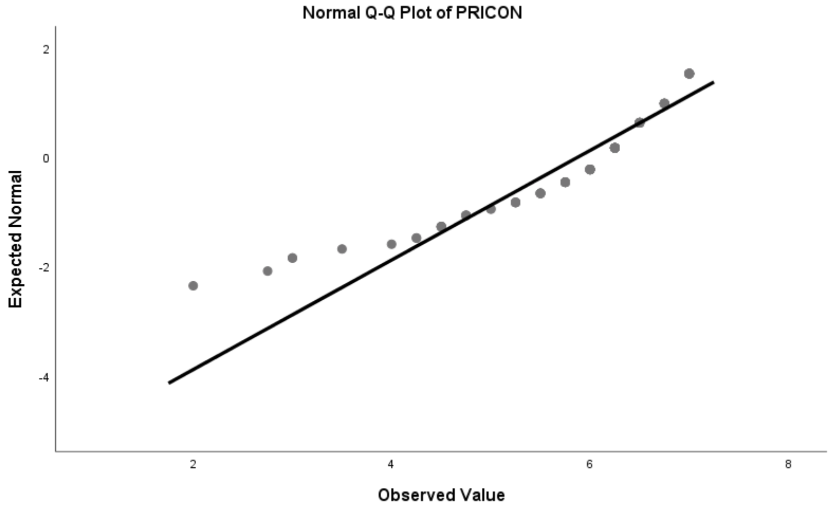

Figure 9. Normal Q-Q plot of PRICON

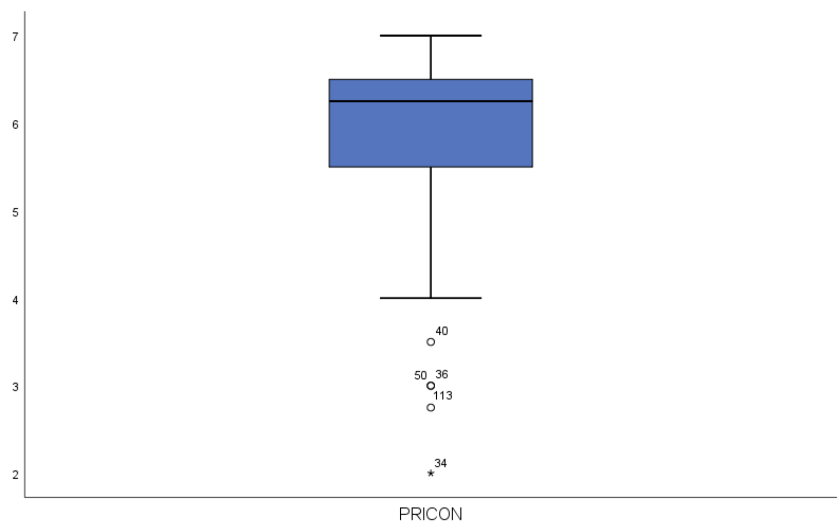

Figure 10. Boxplot of PRICON

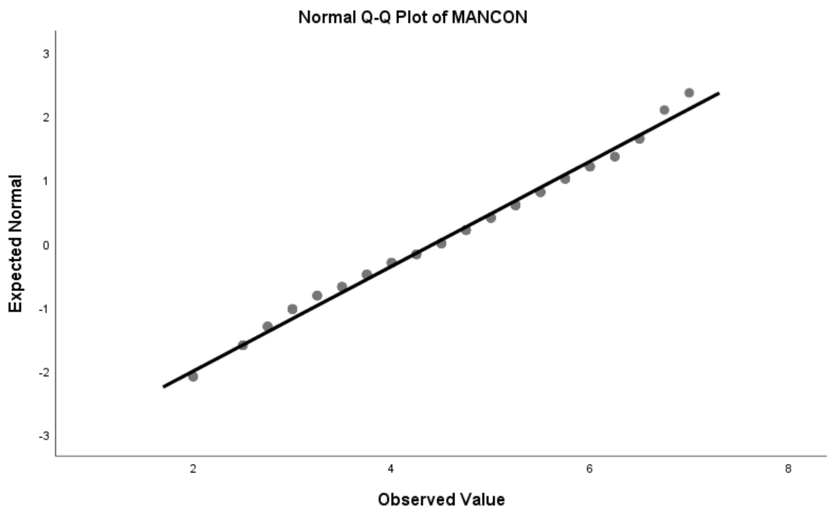

Figure 11. Normal Q-Q plot of MANCON

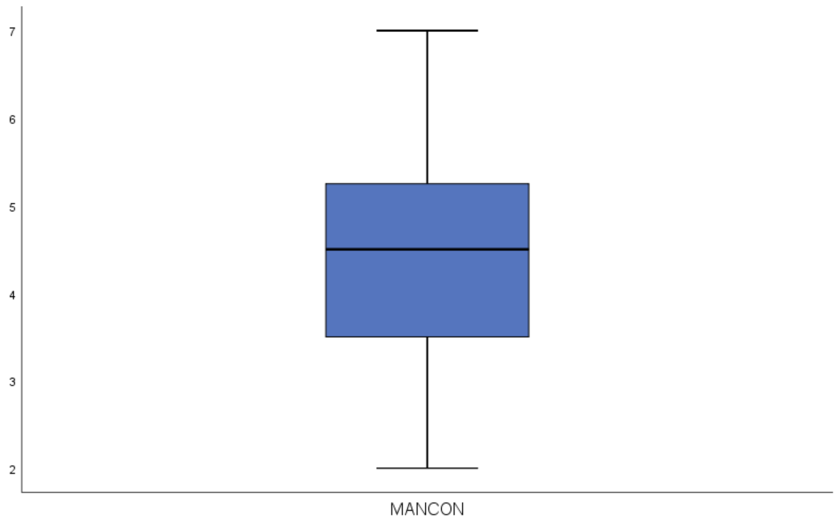

Figure 12. Boxplot of MANCON

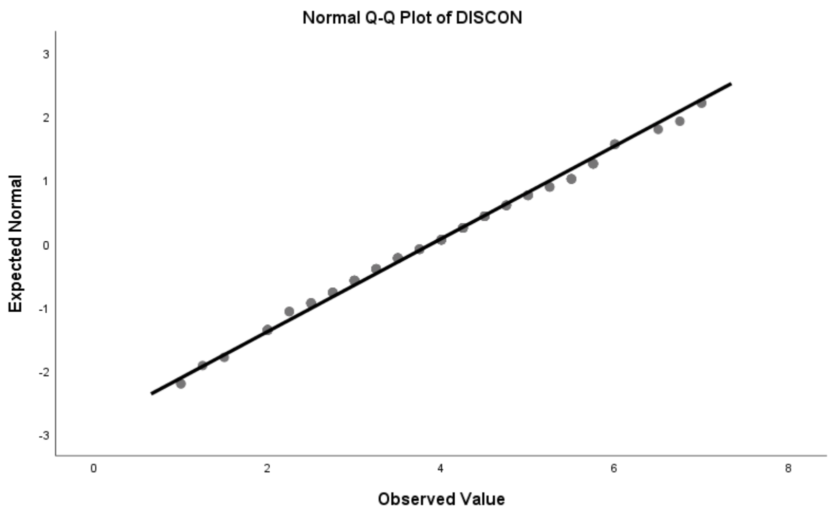

Figure 13. Normal Q-Q plot of DISCON

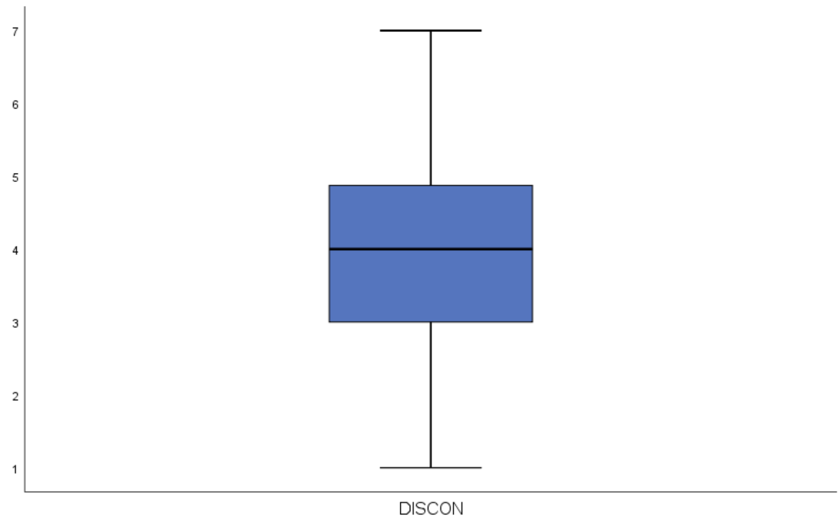

Figure 14. Boxplot of DISCON

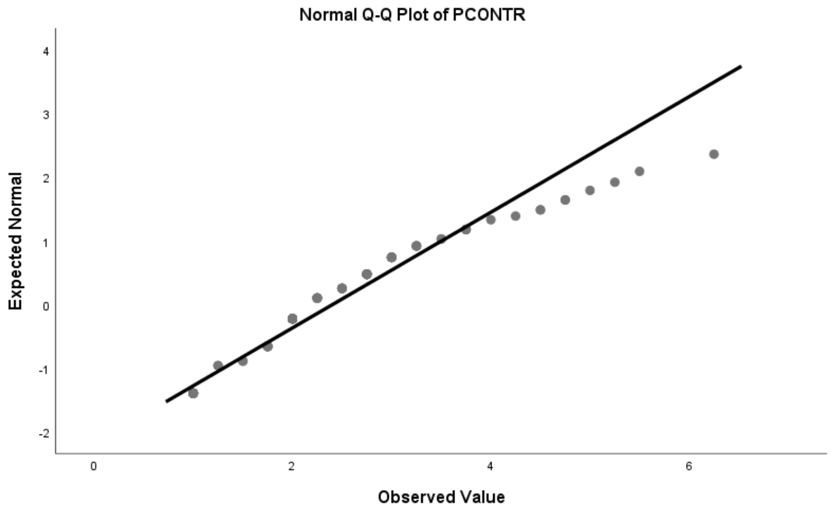

Figure 15. Normal Q-Q plot of PCONTR

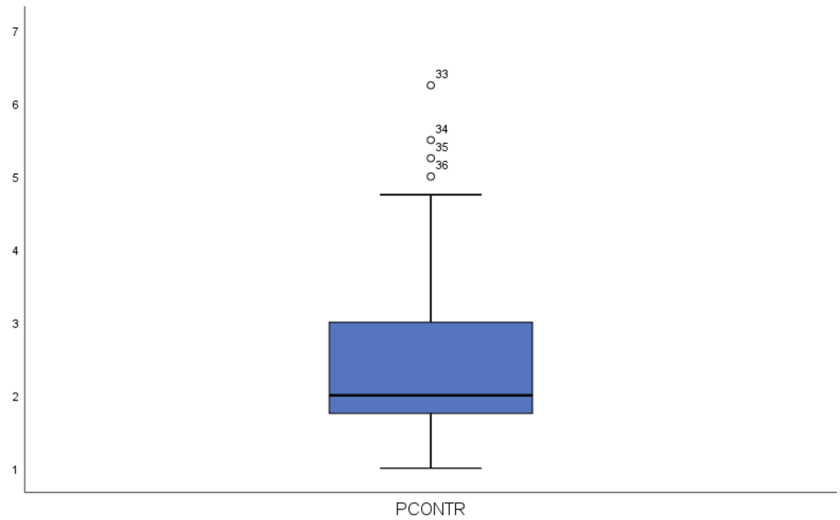

Figure 16. Boxplot of PCONTR

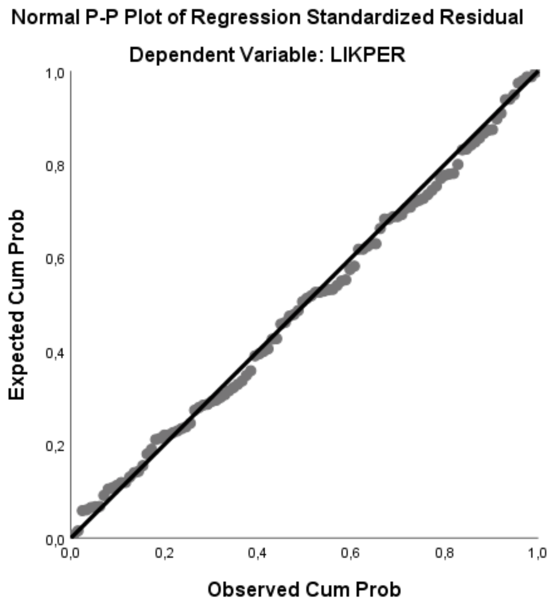

Figure 17. Normal P-P plot of regression standardized residuals.

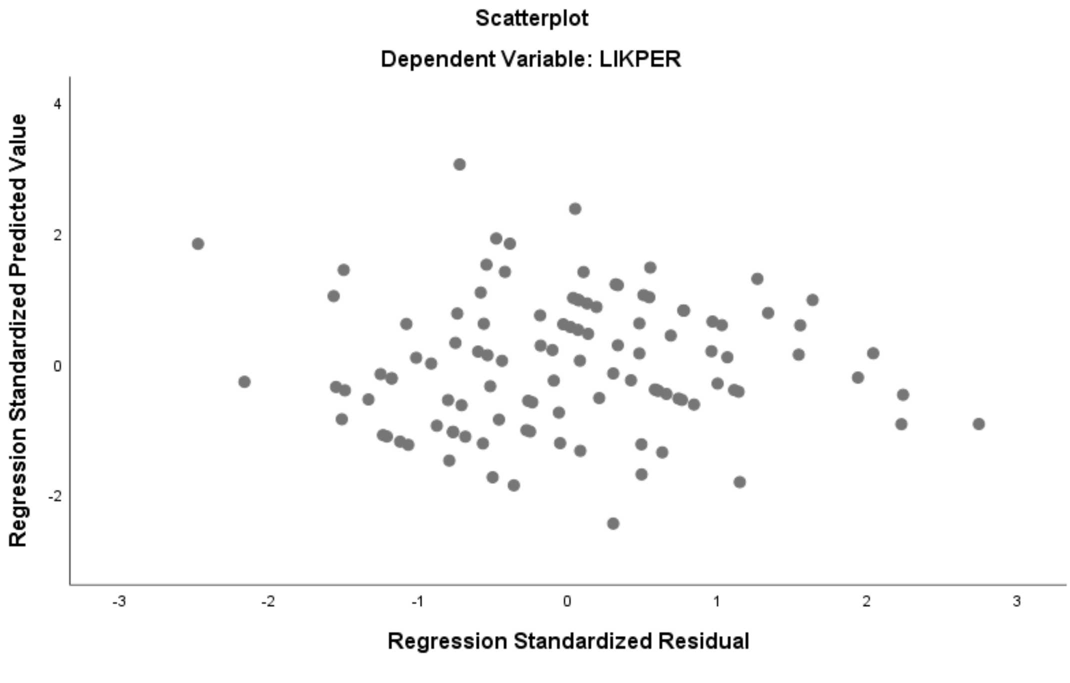

Figure 18. Scatterplot of standardized residuals against standardized predicted values.