The General Data Protection Regulation and the Privacy Paradox: an exploratory study

A first insight of the implication of the new law of data protection on the behavior of high educated Dutch consumers

Shailesh Balla*

December, 2017

ABSTRACT

This research into the privacy paradox and the General Data Protection Regulation is an exploratory study. A first step in providing an answer to what extent high educated Dutch consumers change their behavior after the GDPR. Using an experiment, an environment was created which could possible lead to measuring the first implication of a law that has not been enforced yet. The population of the experiment was divided in three groups, namely two control groups and one treatment group. The difference between these groups was that the first control group got information about the current legislation. The second control group did not receive any additional information. The treatment group received information about the new privacy law. There were four hypotheses tested using regression analysis. There was not enough evidence to reject the null hypotheses. At this moment, there is not enough evidence found to conclude that consumers change their behavior after the General Data Protection Regulation.

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Name student: Shailesh Balla
Student ID number: 369613
Master thesis Economics of Management and Organisation

Supervisor: prof. dr. H.D. Webbink
Second assessor: prof. dr. A.J. Dur
Date final version: 17 December 2017

*The author is grateful for the useful comments and guidance provided by prof. dr. H.D. Webbink. The author also thanks KPMG Netherlands and in particular GRC Technology for the internship and their guidance. Any correspondence to the author can be directed to Shaileshballa@gmail.com.

Table of Contents

I. Introduction

Do consumers care about privacy? Consumers state that they care about their privacy, but researchers found that consumers are willing to disclose personal information in exchange for a reward. This phenomenon is called the 'Privacy Paradox' (Barnes, 2006; Taddicken , 2013).

Brown (2001) discovered the first signs of the privacy paradox, using a series of in-depth interviews with online shoppers. He found that individuals were concerned about their privacy being infringed, but they were still willing to give their personal details to online retailers as long as they had something to gain in return. Interviewees were afraid that companies collected too much information about them, but it did not stop them from buying online. Recently, Athey, Catalini, and Tucker (2017) studied the distortions in consumer behavior when it was faced with notice and choice, which may limit the ability of consumers to safeguard their privacy. They used data from a field experiment (N=3108) from the MIT digital currency experiment. Their research found that a small incentive might explain the privacy paradox: consumers state that they care about their privacy, but if they have a small incentive, they relinquish private information quite easily.

In the European Union (EU) every country has their own privacy legislation and companies extract more data from consumers than necessary. To protect the data of consumers, the EU parliament has come up with the General Data Protection Regulation (GDPR) which has been approved on 14 April 2016 and will be enforced from 25 May 2018 onwards. This law is designed to harmonize data privacy laws across Europe and protect the personal data of all EU citizens (EU GDPR, 2016a). There are five key changes: first, the territorial scope increased. Every company that trades with the EU has to obey the GDPR. Second, the penalties has been increased drastically: four percent of the annual global turnover or €20 million (whichever is greater). Third, consumers have to give explicit consent for the data processing through an easy and accessible form. A company cannot use illegible disclaimers anymore. At last, a consumer has the right to be forgotten. A company has to erase the data of consumers on request (EU GDPR, 2016b).

This research studies the privacy paradox in relation to the GDPR. The GDPR is an opportunity to research further into the privacy paradox. As far as we are concerned, we are the first to write about disclosing personal information and the GDPR. This research focusses on the willingness to disclose personal information under the new privacy law. The intuition behind this is that consumers might trust the data controller more after the implementation of the GDPR, which could lead to consumers disclosing more personal information. The GDPR is a new regulation and the implications of this law are yet unknown. It could be that companies do not obey the privacy

law, which could lead to a more dangerous situation than before the GDPR. Hence, privacy paradox. The research question is:

"To what extent do high educated Dutch consumers change their behavior after the GDPR?"

The scientific relevance of this thesis is that there is not much literature about the GDPR. Also, the privacy paradox has not been studied much in the Netherlands. This is one of the first, if not the first, study into the privacy paradox and the GDPR. This research is meant as an exploratory study, the first step into the research of consumer's privacy and the GDPR. Further researchers can built further on this study.

The societal relevance of this research is that the GDPR is a new law and it is still unknown what the impact of this law is going to be. We attended a seminar of Project Privacy at TQ, a tech-hub in Amsterdam, where they elaborated the new privacy law. They stated that big firms probably will be investigated, but small firms will be fine unless someone complains about them. Consumers should trust that every company processes their personal data well, regardless of the size. It is possible that consumers will disclose more personal data, because they have more trust in companies protecting it. This could be a more dangerous environment, as companies may not comply with the GDPR. This research could be a first insight for consumers to be careful with disclosing their personal data until they know the real implications of the GDPR.

The new privacy law has not taken into place. There exists no data of the state after the enforcement of the GDPR. To research the first implications of this law, this research is conducted as an experiment. The respondents were divided into three groups: two control groups and one treatment group. With the use of survey software it is possible to randomize between three different surveys with equal probability. All three surveys should be represented evenly and not differ in characteristics. All groups see a random text of the University of Leiden about privacy. The first control group receives additional information about the current legislation in the Netherlands. The second control group does not obtain extra information. The treatment group receives additional information about the current legislation and the new privacy law. The first control group fills the questionnaire based on their preferences of the current legislation. The second control group should fill in based on their own preferences. The treatment group should fill the questionnaire in as if the GDPR is the privacy law. With this research design, it is possible to observe the first signs of the implications of the GDPR on the behavior of consumers.

This research consists of nine sections. First, the theoretical framework will be described and more elaboration of the experiment. Thereafter is the data section, research methodology and the results of the experiment. In the bibliography there is an overview of the literature used in this research. Final, the appendices provides extra information of this research.

II. Theoretical Framework

This section describes how the General Data Protection Regulation came to existence by describing the process of the privacy regulation of the European Union. In the literature review, previous studies about privacy disclosure behavior are examined and how this led to the privacy paradox. Furthermore, the mechanism of a person's disclosure behavior is described using a conceptual model. At last, this section elaborates how this relates to this study about the possible change in behavior of consumers and the GDPR.

Privacy regulation

The General Data Protection Regulation has been passed on 14 April 2016 after four years of debating. The new regulation will be enforced from 25 May 2018. Before the new privacy regulation takes into place, there were some important regulatory events that led eventually to the GDPR.

The regulation of privacy started with the publication of the guidelines surrounding data privacy from the organization for Economic Co-operation and Development (OECD). On 23 September 1980, they published eight principles for the processing of personal data. These eight principles were: collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation and accountability principle (OECD, 1980). These principles were conceived to protect the personal data of consumers. Even though this set of principles is 27 years old, it is still relevant today. Many European Countries made their national privacy regulation based on the guideline of the OECD (EU GDPR, 2016a).

Although many European countries made their privacy regulated based on the OECD guidelines, they were still guidelines. It was not binding, which led to a great variety of data protection amongst different European Member states. As an answer to this problem, the European Union implemented the 'Data Protection Directive 95/46/EC' on 24 October 1995, which also was based on the guidelines of the OECD. It was conceived to harmonize the data protection laws of the European Union member states. The main change was the establishment of independent public authorities (DPA's), which were brought to life to supervise the application and the enforcement of the new directive (European Union, 1995).

Despite the attempt of the European Union to harmonize the privacy regulation, directive 95/46/EC was not sufficient. The member states could make their own regulation as long as it was in line with the directive. The directive left room for interpretation, which did not overcome the problem of the great variety of data protection amongst the European Member states. This and the rapid change of the data landscape led to the initial proposal for an updated data

protection regulation on 25 January 2012, the General Data Protection Regulation (EU GDPR, 2016a).

After four years of debating the European Union parliament approved the new regulation on 16 April 2016. After a two-year transition period the GDPR will be enforced on 25 May 2018 (EU GDPR, 2016a).

General Data Protection Regulation

The key difference between the GDPR and the Directive 95/46/EC is that the GDPR is a regulation instead of a directive. This means that it is an enforceable law in all member states. The new privacy directive will mainly influence firms and governments, but it also gives more control for consumers. The aim of the GDPR is to protect the EU citizens from privacy and data breaches.

Before the key changes are discussed, first the definitions as given in the official report of the European Union.

"'Personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (European Union, 2016, p. 33)

"'Processing' means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means such as collection, recording, organization, structuring, storage, adaption or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction' (European Union, 2016, p. 33).

"controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law' (European Union, 2016, p. 33).

Territorial scope

"This Regulation applies to the processing of personal data in the context of the activities of an establishment of controller or a processor in the Union, regardless of whether the processing takes place in the Union or not" (article 3). In comparison to the previous directive, the GDPR will not only affect EU organizations, but also organizations all over the world if they offer services or products to EU citizens. If an organization is not established in the Union (e.g. without a physical

head quarter), it should appoint a representative of the company inside the European Union (article 27).

Penalties

In agreement with the directive 95/46/EC of 1995 every EU member state has his own supervisory authority. If an organization does not obey the GDPR, the supervisory authority has the right to fine. There exists two categories of fines, namely low and high. The low fine has a maximum of €10.000.000,- or up to 2 percent of the worldwide annual turnover of the preceding financial year, whichever is greater. Whereas the high fine has a maximum of €20.000.000,- or up to 4 percent of the worldwide annual turnover of the preceding financial year, whichever is greater (article 83). Under current Dutch legislation, the 'Authority Personal Data[1]' can fine up to €900.000 (Autoriteit Persoonsgegevens, 2016a).

Breach notification

Under the GDPR organizations are obliged to notify the supervisory authority in case of personal data breach within 72 hours after they become aware of the breach. Unless it is unlikely that the breach will result in a risk to the rights and freedoms of natural persons (article 33). From 1 January 2016 organizations in the Netherlands are already obliged to notify the 'Authority Personal Data' in case of a data breach. In some cases they are also obliged to notify the individuals whose personal data has been breached (Autoriteit Persoonsgegevens, 2016b).

Rights of the data subject

The GDPR has been developed to give more right to individuals about their personal data. This is written in chapter III of the GDPR, which is divided in five sections. The first section states that organizations should provide a transparent, intelligible and easily accessible form. They should use clear and plain language to ask consent from the data subject about processing their data (article 12). Second, on request the data controller is obliged to provide information of the data he processes of the data subject within one month, with an extension to three months (article 13). Third, the data subject has the right to be forgotten. On request the data controller has the obligation to erase data concerning a data subject (article 17). They also have the obligation to send the data to another data controller on request (right to portability) (article 20). Fourth, the data subject has the right to object, he can easily withdraw his consent (article 21). The last section explains the privacy by design. Controllers should only process data needed for their business. There are some restrictions of the previous articles, e.g. for national security (article 23).

[1] Dutch public authority for enforcing the legislation of personal data

Data Protection Officer

Public authorities and large organizations who processes data on large scale has to appoint a Data Protection officer (DPO). The DPO is responsible for the internal record keeping requirements and should be appointed on the basis of his professional qualities and knowledge on data protection law. He is also the person who contacts the DPA if there is a breach and has to report directly to the highest level of management (article 38).

Literature study

The literature study elaborates the research of the privacy paradox and how it came to an existence. It also elaborates about studies of the trust in privacy legislation.

Privacy paradox

There are many studies about the relationship between consumers' privacy concerns and their willingness to disclose personal information. One of the first studies on general privacy concerns was from Westin (1967). He wrote a book about privacy and freedom, where he evaluated the conflict between individual privacy rights and surveillance in modern society. He believed that there was need for legislation to safeguard the rights of the people. However, he advocates a balanced position, which meant that limited use in cases of national security and major crimes was allowed (Westin, 1967). The earliest privacy index was the "General Privacy Concern Index," which he developed in a study in 1990. Westin, Harris and Associates (1990) divided consumers into three different categories: the privacy fundamentalist who are generally distrustful of organization that ask them personal information. The Pragmatic who weigh the benefits and costs of providing personal information. The Unconcerned who are trustful of organization collecting their personal information (Westin, Harris, & Associates, 1990).

The "General Privacy Concern Index," also known as the West-Harris privacy index were useful indications of the privacy concerns of consumers. However, it could not provide information on how individual information disclosure decisions are affected by privacy concerns (Motiwalla, Xiaobai, & Xiaoping, 2014). Which led to research focusing on individual disclosure decisions under risk uncertainty (Kahneman & Tversky, 1984). Focusing on how privacy concerns affect individual disclosure behavior led to the invention of the privacy calculus model. This model is based on the behavior of individuals who make privacy decisions according to their evaluations of the risks and benefits induced by the information disclosure behavior (Culnan & Bies, 2003). This model suggests that people are less willing to disclose personal information if the perceived risks are higher and vice versa (Dinev & Hart, 2006; Belanger & Crossler, 2011; Li, Sarathy, & Xu, 2010; Krasnova & Veltri, 2010). Privacy researcher came to the consensus that consumers can have both benefits and risks when disclosing their personal information (Culnan & Bies, 2003; Krohn, Luo, & Hsu, 2002).

While the privacy calculus was a new model to research privacy, many researchers found an inconsistence between the information disclosure behavior and privacy concerns. This led to the invention of the privacy paradox. Privacy paradox research (Athey et al., 2017; Barnes, 2006; Taddicken, 2013; Gross & Acquisti, 2005) found that consumers say that they care about privacy, but are easily willing to disclose personal information when they are incentivized to. One of the first studies about privacy paradox was from Brown (2001). Using a series of in-depth interviews with online shoppers, Brown found that individuals were concerned about their privacy being infringed, but they were still willing to give their personal details to online retailers as long as they had something to gain in return. Interviewees were afraid that companies collected to much information about them, but it did not stop them from buying online. Taddicken (2014) found, using a sample of German Internet users (N=2,739), that people are concerned about their privacy, but it does hardly impact their self-disclosure. Where Smith, Dinev, and Xu (2011) goes even further and assumes that consumers with higher risk attitudes will nonetheless disclose their personal information for relatively low benefits in high-risk situations and vice versa.

Previous literature (Hui, Teo, & Lee, 2007; Beresford, Kübler, & Preibusch, 2012; Hann, Hui, Lee, & PNG, 2007; Grossklags & Acquisti, 2007) studied the monetary value of consumers' willingness to disclose information with field- and lab experiments. They found, using proxies, that consumers are willing to disclose personal information in exchange for money. They studied the monetary value, which was still the observed willingness of a price instead of their actual behavior.

Although the privacy calculus model and privacy paradox studies could examine how individual information disclosure decisions are affected by privacy concerns, it still had limitations. This comes from the fact that they examine the consumer willingness to provide information, but not their actual disclosure behavior. One can say that they are concerned about privacy, but handle quite differently. The models captures their intentions and not their real behavior (Xu, Teo, Tan, & Agarwal, 2009; Smith et al., 2011). To overcome this limitation, researchers observed actual consumer behavior by trading personal information in exchange for money (Motiwalla et al., 2014; Huberman, Adar, & Fine, 2005; Athey et al., 2017). This limitation can be overcome by trading personal information in exchange for money. In this research we do not trade personal information in exchange for money because of the lack of financial resources.

Trust in privacy legislation

Between 1978 and 2004, Alan Westin conducted over more than 30 surveys for which he created the 'Consumer Privacy Index. Unfortunately, most of the early studies were distributed in paper and therefore not readable anymore. Some of the later papers are available for purchase. Luckily,

Kumaraguru and Cranor (2005) studied the methodology and results of six privacy indexes, which is used to describe the trust in privacy legislation over the years.

The first study Kumaraguru and Cranor (2005) discusses is the Westin (1991) study. In this survey he asked about the trust in the privacy legislation through the question: "my privacy rights as a consumer in credit reporting are adequately protected today by law and business practices." 1,255 citizens of the United States participated in this research. The question was measured on the five-point Likert scale. The result of this research was: agree very strongly (37), agree somewhat strongly (34), disagree somewhat strongly (20), disagree very strongly (3) and neither/not sure (4). The numbers in parenthesis are percentages. The majority of the sample did not think that their privacy rights were adequately protected by the privacy legislation.

The second study was the study of 1994 " (Westin, Harris, & Associates, 1991) where they measured the trust in the privacy legislation through: "Government can generally be trusted to look after our interests." In this research 1,005 United States citizens participated and the results were: agree strongly (5%), agree somewhat (15%), disagree somewhat (28%) and disagree strongly (52%).

In 1996 Westin rephrased the question to: "Which of these choices do you think is best for the U.S?" 1,005 United States citizens could choose between three options: creating a federal government privacy commission, using the present system to protect consumer privacy or neither. The result of this question was respectively 28, 67 and 5 percent. Which meant that the majority of United States citizens did not want to change the present system.

In 1999 and 2000, they asked: "Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today." (Westin & Harris, 1999; Westin & Opinion Research Corporation, 2000). In 1999 59 percent strongly/somewhat agree against 51 percent in 2000. Logically, 38 percent in 1999 and 49 percent in 2000 strongly/somewhat disagree. In the 2001 study, Westin (2001) asked the same question and found: strongly disagree (4%), somewhat disagree (34%), somewhat agree (45%) and strongly agree (18%). Which combines to 62 percent strongly/somewhat agree. In 2003, 47 percent agreed with the reasonable level of protection of existing law (Westin, 2003).

The research methodology was changed by Westin every year, which is why it is difficult to observe a reliable measurement of the trend of the trust in privacy legislation. There are more studies about the trust in privacy legislation, but Westin is one of the few who reported the value of trust in privacy laws. In most papers read, there was no detailed information about the privacy concerns of privacy legislation.

Conceptual model

In previous privacy literature, studies used figure 1 as the mechanism how person's disclosure behavior works (Bart, Shankar, Sultan, & Urban, 2005; Schoenbachler & Gordon, 2002). Both risk and trust influences the behavioral intention to disclose. The individual then chose what personal information he was willing to disclose. Norberg, Horne, and Horne (2007) were the first who altered this model into a new mechanism. Behavioral intention (to disclose) was influenced by the risk of disclosing information. The real disclosure behavior is influenced by the trust in the receiver of the information (as seen in figure 2).

With the use of undergraduate students, they test their modified model in a classroom setting. The students got two different scenarios: a low and high-trust scenarios. After reading these scenario's they had to fill in which personal information they are willing to provide personal information to a company in exchange for $20 (e.g. personal health history, monthly income, dining preferences). After these scenarios they were asked how trustworthy, honest and sincere these company were. They also were asked about how risky it was to provide personal information to the company. They see their study as an exploratory study to investigate the privacy paradox. Using the new model, they found evidence for the privacy paradox where the students based their disclosure behavior on the level of trust (Norberg et al., 2007).

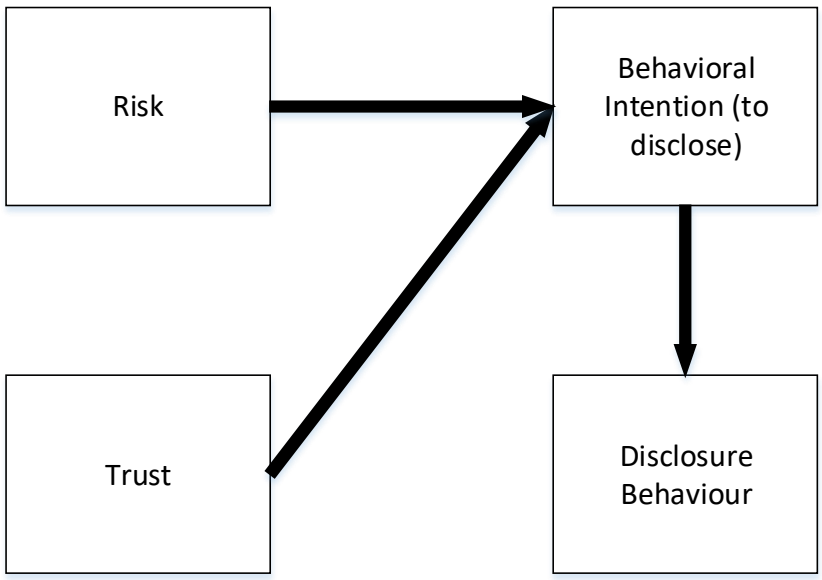

Figure 1: Conceptual model of disclosure behavior based on previous research (Norberg et al., 2007)

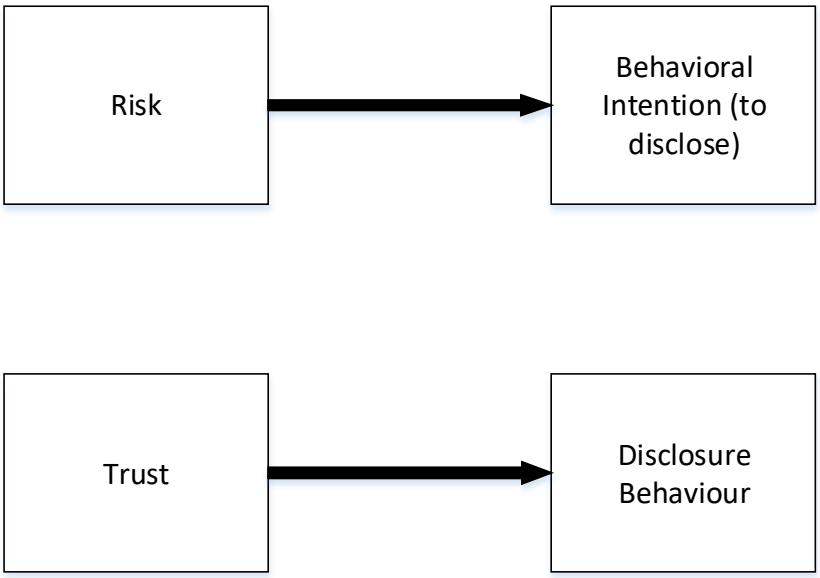

Figure 2: Conceptual model of disclosing behavior - Privacy Paradox (Norberg et al., 2007)

This research

The privacy paradox is defined as one stating that they are concerned about their privacy, but are easily willing to disclose personal information when incentivized. In this research, the focus lies on the disclosure behavior model of Norberg et al. (2007), where risk influences behavioral intention (to disclose) and trust influences the real disclosure behavior. It only focusses on the level of trust. There are three levels of trust: control group 1 who trust the current privacy law, control group 2 who trust their own preferences and the treatment group who trust the new privacy law.

The new privacy regulation is made to protect the processing of personal data of consumers. It might lead to more trust in the privacy legislation, which could lead to consumers more willing to disclose personal information. The GDPR has not been enforced yet, so the implication of this legislation are unknown. It is possible that companies will not comply with the new privacy law, leading to more risk. If a consumer has more trust in the GDPR, it could lead to a more dangerous situation if companies do not comply. This is the privacy paradox researched in this research.

Hypotheses

Using four hypotheses, the research question: "To what extent do high educated Dutch consumers change their behavior after the GDPR," will be answered. These hypotheses will be tested using regression analysis and should give enough evidence to accept or reject the alternative hypothesis, leading to answering the research question.

The first hypothesis is about the change in willingness to disclose information after the GDPR. Norberg et al (2007) found that students were more willing to disclose information when they have more trust in an organization. The new privacy legislation is made up to protect the processing of personal data of consumers, so consumers should have more trust in this legislation. Therefore they might be more willing to disclose personal information. The hypothesis is as follows:

H1a: Consumers are more willing to disclose personal information after the enforcement of the GDPR.

The second hypothesis is about the trust in companies. Companies should obey the new privacy regulation. The GDPR holds for all companies who processes personal data of EU citizens, no matter where they are based. This hypothesis is to test if the consumers' trust in the compliance by companies increases after the enforcement of the GDPR. Consumers should trust companies in handling their personal data well, otherwise they might not be willing to disclose personal information. The second hypothesis is as follows:

H2a: consumers have more trust in the compliance of the privacy law by companies after the enforcement of the GDPR

The third hypothesis is about the trust of consumers in the enforcement of privacy law by the authorities. Dutch consumers should trust the 'Autoriteit Persoonsgegevens' in enforcing the law and fining a company if they do not obey the GDPR. If consumers do not trust authorities will enforce the privacy law, then they might be less willing to disclose personal information. So under the new privacy legislation, there should be more trust in the enforcement of the authority. The third hypothesis is as follows:

H3a: Consumers have more trust in the enforcement of the privacy law by the authorities after the enforcement of the GDPR.

The fourth hypothesis is about the overall trust in the privacy legislation. The GDPR is adopted to protect the processing of personal information of the consumers of the European Union. Consumers should have more trust in the new privacy law for it to be effective. The fourth hypothesis is as follows:

H4a: Consumers have more trust in the privacy law after the enforcement of the GDPR

With these four hypotheses, the research question "to what extent do high educated Dutch consumers change their behavior after the GDPR?" is tried to be answered.

III. The Experiment

The GDPR has not been enforced yet, so there is no data available about the behavior of consumers after the GDPR. To overcome this, the research is conducted as an experiment with two control groups and one treatment group. The respondents did not know that they were part of an experiment.

With the use of survey software, it is possible to randomize with equal probability between these three surveys. All surveys should be presented evenly and should not differ in characteristics. One caveat of this design is when a person does not complete the survey, it still counts as a view. This might lead to a bias if one particular survey was not filled in completely. This results in less observations for that particular group and maybe biasing the results. The groups will be tested for the hypothesis that they are from the same population.

The three groups received the same questionnaire, which can be found further in this research, but differ in the information they received at the start of the survey. All groups received a random text in the beginning from some research of the University of Leiden:

"Netherlands in the leading group on privacy protection." Research from Leiden University has shown that the Netherlands is doing above average when it comes to the protection of personal data. Germany is leading, countries like Italy and Romania are lagging behind. They compared eight European countries to various privacy aspects, such as government policy, legislation, and supervision and enforcement.

The first control group received additional information about the current legislation in the Netherlands:

Under the privacy law, organizations may only process personal data that is necessary for the organization. The person whose personal data are processed must at least be aware of the identity of the organization and the purpose of the data processing. In the Netherlands, the 'Autoriteit Persoonsgegevens' is responsible for enforcing this law. When an organization does not comply with the privacy law, it can impose a fine up to €820.000.

The second control group did not receive additional information. The treatment group received information about current privacy law and the GDPR:

*From **May 25, 2018** the new privacy law takes into place (General Data Protection Regulation). The most important changes are:*

- *The law applies to the entire European Union and every country that trades with the EU.*
- *A consumer must explicitly give permission for a specific purpose of the personal data*
- *Right to access: on request an organization is obliged to provide a copy of the data they have from a person*
- *Right to be Forgotten: companies are obliged to remove personal data on request*
- *Data portability: a person has the right to transfer his personal data*
- *Privacy by design: companies are obliged to apply the default option in such a way that it protects the privacy of consumers.*

In the Netherlands, the 'Autoriteit Persoonsgegevens' is responsible for the enforcement of this law. In the current situation she may impose fines up to €820.000. From May 25, 2018, this fine will be increased to €20 million or 4 percent of the worldwide turnover (whichever is greater).

The first control group receives information about the current privacy law. That should influence them to fill in the questionnaire based on their beliefs of the current privacy law. The second control group receives no additional information. They should fill in the questionnaire based on their prior beliefs of the privacy law. Both control groups should have similar response, namely of the current legislation.

The treatment group receives additional information about the new privacy law. They should fill in the questionnaire as if the GDPR has taken into place. To verify this, the treatment group received an additional question: "You have to fill in the following questions as if the 'General Data Protection Regulation' has already been enforced (thus after 25 May 2018)." 125 respondents understood that they should fill the questionnaire as if the GDPR has taken into place, against 3 respondents who did not understood. Below there is a table of the questions asked.

Table 1: Questionnaire used in this research

	Question	Scale
Q3	I am familiar with the General Data Protection Regulation?	Yes/No
Q8	You want to start a company and you ask company "SB" for some advice. How more personal data company "SB" has, how better the advice. I am willing to provide the following data:	
	- E-mail	Yes/No
	- Medical data	Yes/No
	- Monthly purchases	Yes/No
	- Address	Yes/No
	- Phone number	Yes/No
	- Credit card and cvc code	Yes/No
	- Date of Birth	Yes/No
	- Income	Yes/No
	- Gender	Yes/No
	- Political preferences	Yes/No
	- Name	Yes/No
	- Religious belief	Yes/No
Q9	I trust that companies will handle my personal data carefully	7 point Likert
Q10	I trust that the data protection authority will enforce the privacy law	7 point Likert
Q11	Countries of the European Union satisfy the privacy law on average	7 point Likert
Q12	I find the fine for not complying with the privacy law sufficient	7 point Likert
Q13	The privacy law provides a reasonable level of protection for my personal data	7 point Likert
Q14	I am willing to share my data if I know the data collector or website	7 point Likert
Q15	Sharing my data will help me access better products and services	7 point Likert
Q16	Sharing my data will make me more vulnerable to identity theft	7 point Likert
Q17	How old are you?	
Q18	What is your gender?	Male/Female
Q19	Are you married?	Yes/No
Q20	How many children do you have?	0, 1, 2, 3, 4 or more
Q21	What is your highest level of education?	Primary school, MAVO, HAVO, VWO, MBO, HBO, Wo Bachelor and WO Master
Q22	What is your occupation?	Study, part time work, full time work, job seeking and others
Q23	What is your income?	< €10.000, €10.000 - €20.000, €20.000 - €30.000, €30.000 - €40.000, €40.000 - €50.000 and €50.000 <
Q24	What is your country of origin?	Netherlands, Surinam, Dutch Antilles, Turkey, Morocco and others

IV. Data

Section four describes how the data was collected. It also describes which variables are used to test he hypotheses and there is an overview of the descriptive statistics. At last, there is a test of equality of population, to test if the three groups are from the same population.

Data collection method

The data used in this research is survey data which was collected in the period of 18 October until 4 November 2017 from Dutch citizens. This research focuses on the behavior of Dutch citizens. To reduce misunderstandings, the questionnaire was held in Dutch.

At first, this research focused on the whole Dutch population. Using the formula of Barlett, Kotrlik, and Higgins (2001) the population sample was calculated, which were 384 observations. Then the survey was spread through Facebook, LinkedIn, WhatsApp and emailing over a hundred KPMG colleagues. This led to a total response of 489 individuals. A first look at the data showed that the education level of the dataset was highly skewed to high educated citizens. Approximately 78 percent of the respondents had finished at least their study at a University of Applied Sciences. That is why this research focusses on high educated Dutch citizen. The observations of individuals who have obtained a degree below the degree of the University of Applied Sciences, led to a sample of 396 respondents. Three observations with the age of zero were deleted. The total respondents of this research is 393. If the total population is above one million, the population sample does not change. Calculation of the population sample is provided in the appendices.

Variables

The questions asked in the survey are based on literature and sometimes slightly modified for this research. The order in which the respondent could fill in the questions were randomized. This means that every respondent saw the questions in a different order. The intuition behind this is that with the use of randomization a framing bias can be overcome.

To test the hypotheses, four different questions are used as dependent variables. The first hypothesis is about the willingness to disclose personal information. Following the research methodology of Norberg et al. (2007), a small case was used where the respondent could fill in if he is willing to disclose personal information to company 'SB' to receive a tailor-made offer. In the case is stated explicitly that the tailor-made offer will be better, if company 'SB' has more information of the respondent. The respondent could choose, using a yes-no question, if he is willing to disclose the following information: e-mail, medical data, monthly purchases, address, phone number, credit card number and cvc code, date of birth, income, gender, political

preference, name and religious belief. The order of the items were randomized, so a framing bias could be overcome.

In order to test the other three hypotheses, the questions are based on the research of Motiwalla et al. (2014). Their questions are based on the Westin's consumer privacy index and measured on the seven-points Likert scale ranging from strongly disagree, disagree, disagree somewhat, undecided, agree somewhat, agree and strongly agree. Whereas '1' is strongly disagree and '7' is strongly agree. The literature of the Likert scale (Likert, 1932) cannot agree on which scale is best, the five- or seven point scale. A few researches reported higher reliabilities for five-point scales (Jenkins & Taber, 1977; McKelvie, 1978). Whereas some found the seven point scale more suitable (Finstad, 2010; Preston & Colman, 2000). Overall, John (2010) found that Likert scales becomes significantly less accurate when the number of scale points drops below five or above seven. The seven points Likert scale is used because Likert (1932) and other researchers recommend that it is best to use a scale as wide as possible. It is always possible to categorize the answers (Allen & Seaman, 2007).

To test if the seven point Likert scale was reliable, the Cronbach's Alpha was calculated (1951). The alpha says something about the internal consistency of the scale used. When the alpha is above the 80 percent level, researches considers this as a highly reliable scale. When the alpha is below the 50 percent level, it is considered as a highly unreliable scale. In social sciences, a Cronbach's alpha between 70 and 80 percent is considered as acceptable (Tilburg University, 2017). The alpha was below 70 percent. Deleting three questions, which were not used in this research, led to an alpha of approximately 71 percent. This made the seven point Likert scale an acceptable measurement for this research. The calculation of the Cronbach's Alpha can be found in the appendices.

The independent variable is the group in which the respondent is placed. This are three dummy variables, where for each dummy variable '1' means that they are in that particular group and '0' otherwise.

Besides the dependent and independent variables, there are also background variables to get more information about the characteristics of the respondents. They were used to test if the groups have similar characteristics and as controls in the models. The background variables are: age, gender, marriage, children, level of education, primary occupation, gross annual income and the country of origin.

The respondents could fill in their age in numbers. This is modified to an ordinal scale for research purposes. The scale ranges from: <20, 20-30, 30-40, 40-50, 50-60 to 60<.

For the country of origin, there were six options, namely Netherlands, Turkey, Morocco, Surinam, Dutch Antilles and others. These are the largest population groups in the Netherlands (Centraal Bureau voor de Statistiek, 2016). There are much more nationalities in the Netherlands, but they were combined to these six categories as this represents most of the Dutch population.

Descriptive Statistics

In total there are 393 respondents in the experiment. The first control group has 137 respondents. Both the second control group and treatment group had 128 respondents. The survey was randomized with equally probability. The surveys should have the same number of respondents. This means that the first control group was finished the most. Respondents of the second control group and treatment group stopped prematurely with the questionnaire.

The proportion males to females was 207 (52.67%) to 186 (47.33%). The majority (65%) of the respondents did not hear of the GDPR before.

The majority of the respondents originated from the Netherlands (84.73%). The second largest group was from Surinam (9.92%). Three respondents were originated from Morocco (0.76%) and one from Turkey (0.25%). There were seventeen respondents (4.33%) from other countries.

The descriptive statistics of the groups are given on the next page. This corresponds with the questionnaire of the experiment. It should be interpreted as: 97 percent of the respondents who were in the first control group, are willing to disclose their e-mail to company “SB”. For the questions measured on the seven point Likert scale, it should be interpret as: the mean of the trust in companies of the control group is 4.37.

Table 2: Descriptive statistics all groups

		Control group I		Control group II		Treatment group			
		N = 137		N = 128		N = 128			
	Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Min	Max
Q3	I am familiar with the GDPR	0,38	0,49	0,33	0,47	0,35	0,48	0	1
Q8	Willingness to share:								
	- E-mail	0,97	0,17	0,99	0,09	0,96	0,19	0	1
	- Medical data	0,07	0,25	0,13	0,33	0,13	0,33	0	1
	- Monthly purchases	0,20	0,40	0,27	0,44	0,27	0,45	0	1
	- Address	0,76	0,43	0,74	0,44	0,70	0,46	0	1
	- Phone number	0,78	0,42	0,83	0,38	0,82	0,39	0	1
	- Credit card and cvc code	0,04	0,21	0,06	0,24	0,03	0,17	0	1
	- Date of Birth	0,86	0,35	0,91	0,29	0,83	0,38	0	1
	- Income	0,32	0,47	0,34	0,48	0,35	0,48	0	1
	- Gender	0,91	0,28	0,95	0,21	0,89	0,31	0	1
	- Political preferences	0,31	0,46	0,27	0,45	0,27	0,45	0	1
	- Name	0,93	0,26	0,91	0,28	0,86	0,35	0	1
	- Religious belief	0,43	0,50	0,38	0,49	0,38	0,49	0	1
Q9	I trust that companies will handle my personal data carefully	4,37	1,69	4,09	1,73	4,04	1,63	1	7
Q10	I trust that the data protection authority will enforce the privacy law	5,12	1,45	4,93	1,56	4,95	1,44	1	7
Q11	Countries of the European Union satisfy the privacy law on average	4,03	1,22	4,08	1,32	3,74	1,35	1	7
Q12	I find the fine for not complying with the privacy law sufficient	4,15	1,71	3,92	1,32	4,75	1,53	1	7
Q13	The privacy law provides a reasonable level of protection for my personal data	4,66	1,26	4,52	1,28	4,88	1,21	1	7
Q14	I am willing to share my data if I know the data collector or website	4,86	1,36	4,95	1,37	4,64	1,59	1	7
Q15	Sharing my data will help me access better products and services	4,42	1,65	4,30	1,74	4,32	1,54	1	7
Q16	Sharing my data will make me more vulnerable to identity theft	5,65	1,04	5,88	1,19	5,78	1,00	1	7
Q17	Age	26,26	9,15	26,53	8,94	27,13	8,70	17	66
Q18	Gender	0,50	0,50	0,52	0,50	0,56	0,50	0	1
Q19	Married	0,12	0,33	0,09	0,28	0,14	0,35	0	1
Q20	Children	0,28	0,80	0,24	0,67	0,37	0,83	0	4
Q21	Level of education	1,16	0,81	1,16	0,81	1,19	0,87	0	2
Q22	occupation	1,95	1,07	2,02	1,13	2,06	1,15	1	5
Q23	Income	3,30	2,53	3,26	2,50	3,46	2,49	1	7
Q24	Ethnicity	1,34	1,12	1,49	1,32	1,23	0,77	1	6

Test of equality of background variables

The respondents were randomly divided into the three groups. It is assumed that they do not differ in characteristics, as the groups were randomly selected. When the groups would differ in characteristics, there might be an overrepresentation of a characteristic, which could possibly lead to a bias of the results. To test whether groups do not differ in characteristics, the Kruskal-Wallis equality-of-population rank test is used (Kruskal & Wallis, 1952). It tests the hypothesis that several samples are from the same population. If the probabilities are below the significance level of 5 percent, it cannot be accepted that the three groups are from the same population. How larger the probability, how more the groups are similar to each other.

Table 3: Kruskal-Wallis test of equality of population of the groups

Variable	All groups	Control group 1 and treatment	Treatment and control group 2	Control groups 1 and 2
Age	0.345	0.181	0.239	0.885
Gender	0.558	0.282	0.531	0.660
Married	0.378	0.692	0.168	0.314
Children	0.324	0.197	0.213	0.945
Level of education	0.909	0.687	0.724	0.965
Occupation	0.810	0.515	0.798	0.712
Income	0.734	0.566	0.453	0.845
Ethnicity	0.515	0.956	0.308	0.346

In the second column of the table above, there is an overview of the probability of the chi-squared with ties. The probabilities are all above the five percent significance level. This means that the hypothesis that the three groups are from the same population cannot be rejected. This is essential for this research. Differences in characteristics could bias the results.

In addition, the hypothesis is also tested for the groups separately, respectively the third, fourth and fifth column. It is possible to observe how large the probabilities are within the three groups. None of the probabilities are below the 5 percent significance level. This means that the hypothesis that they are from the same population cannot be rejected.

V. Methodology

The first hypothesis is going to be tested with the use of a small case about the willingness to disclose personal information to company 'SB'. There are 12 different items of personal information. If the respondent is willing to disclose an item, it gets the value '1'. If not, it gets the value '0'. The other three hypotheses are going to be tested using the questions described in the data section and is measured using the seven-point Likert scale.

To test the hypotheses, different models are going to be estimated using statistical software. The dependent variables are either measured on the binary scale or on the seven-point Likert scale. The appropriate estimation model is the ordered probit or logit. The choice for an ordered probit- or logit model is based on the preference of the researcher. Both estimations give similar results, but are calculated differently and based on different assumptions. In this research, the (ordered) probit model is used. Also, OLS models are estimated as they are easier to interpret and the results should not differ much if the number of observations is large enough. Other regression models are more advanced techniques.

In comparison to the Ordinary Least Squares (OLS), the ordered probit estimation does not assume that the data is linear. This is why the coefficients of the estimations are difficult to interpret. The interpretation of the coefficients of the ordered probit model is: when the coefficient is negative, it means that the probability of being in the highest category decreases if the explanatory variable increases and that the probability of being in the lowest category increases if the explanatory variable increases. Vice versa, if the coefficient is positive, it means that the probability of being in the highest category increases if the explanatory variable increases and that the probability of being in the lowest category decreases if the explanatory variable increases. The magnitude cannot be interpreted, only the sign and the significance. It is also not possible to say any effect of the explanatory variable on the middle categories (García-Gómez, Bago d'Uva, & Riumallo-Herl, 2017)[2].

The ordered probit model normally does not estimate a R^2 as under OLS, but the newest version of Stata a pseudo R^2 is estimated. This statistic does not mean the same as under OLS (e.g. the proportion of variance of the responsible variable explained by the predictors). UCLA Institute for Digital Research and Education (2017) suggests interpreting this statistics should be done with great caution.

The following regression is the one being estimated. *Y* is the dependent variable that varies with the hypothesis. α is the constant and is only going to be estimated with the OLS model. The

[2] Source from Blackboard EUR (not publicly accessible)

ordered probit model does not estimate a constant. β_1 and β_2 are the coefficient of interest, namely of the treatment group and the second control group. X is a vector of all control variables and there will be models with and without controls for background variables. For completeness, ε is the error term of the estimation.

$$Y = \alpha + \beta_1 \cdot Treatment\ Group + \beta_2 \cdot Control\ group\ 2 + X + \varepsilon$$

The first hypothesis is estimated with OLS and probit. The probit results are available on request. The models will be regressed on the items separately. The items will be ordered from information that people are most willing to disclose to information people least willing to disclose. With this setting it is possible to observe if the treatment has effect on information easy or hardly willing to disclose. There might exist a different willingness to disclose.

The last three hypotheses, measured on the seven point Likert scale, are estimated with OLS and ordered probit. The OLS estimation is given, such that the reader can observe the magnitude of the coefficients. The reader should careful interpret the coefficients of the OLS estimation, because the errors (i.e., residuals) violates the homoscedasticity and normality of errors assumptions of OLS, which results in invalid standard errors and hypothesis tests (UCLA Institute for Digital Research and Education, 2017).

The models are estimated with- and without controls for characteristics of the groups. The background variables are categorical. The models with controls have dummies for the options of the background variables.

As the models are estimated, the alternative hypothesis is going to be rejected or accepted. This is based on the significance of the coefficients of interest. When the significance is below the five percent level, the null hypothesis will be rejected. If the significance is above the five percent level, there is not enough evidence to reject the null hypothesis.

Even if the null hypothesis cannot be rejected, the coefficients will be described. If the beta of the treatment group is negative and lower than the beta of the second control group (does not matter if beta of second control group is negative or positive), being in the treatment group gives a lower value than being in both control groups. If the beta of the treatment group is positive and higher than the beta of the second control group, being in the treatment group gives a higher value than being in both control groups. If the beta of the treatment group is positive, but lower than the beta of the second control group, being in the treatment group has a positive effect in comparison with the first control group, but a negative effect with the second control group.

After the hypotheses are tested, the research question is answered as detailed as possible.

VI. Results

This section elaborates the results of the hypotheses tested. Both OLS and probit models are estimated with- and without controls for background variables. The results should be interpreted with caution.

Table 4: Regression estimates of the effect of the treatment on the willingness to disclose specific information

	Without Controls			With controls		
Dependent variable:	Treatment group	Control group 2	(Pseudo) R^2	Treatment group	Control group 2	(Pseudo) R^2
E-mail (97%)	-0,010	0,021	0,007	-0,016	0,013	0,162
	(0,022)	(,016)		(,205)	(,013)	
Name (93%)	-0,068	-0,013	0,010	-0,064	-0,010	0,149
	(0,038)*	(,033)		(,367)*	(,032)	
Gender (91%)	-0,022	0,041	0,009	-0,025	0,027	0,209
	(0,037)	(,031)		(,037)	(,029)	
Date of Birth (86%)	-0,033	0,045	0,009	0,038	0,032	0,213
	(0,045)	(,039)		(,043)	(,036)	
Phone number (78%)	0,039	0,047	0,003	0,038	0,046	0,081
	(0,049)	(,049)		(,050)	(,052)	
Address (76%)	-0,064	-0,017	0,004	-0,081	-0,012	0,099
	(0,055)	(,053)		(,056)	(,054)	
Religious belief (43%)	-0,056	-0,048	0,003	-0,025	-0,035	0,095
	(,060)	(,061)		(,063)	(,062)	
Income (32%)	0,030	0,023	0,001	0,053	0,035	0,162
	(,058)	(,058)		(,059)	(,057)	
Political preferences (31%)	-0,033	-0,033	0,001	-0,016	-0,019	0,065
	(,056)	(,056)		(,059)	(,058)	
Monthly purchases (20%)	0,069	0,061	0,005	0,092	0,055	0,145
	(,053)	(,052)		(,053)	(,052)	
Medical data (7%)	0,059	0,059	0,009	0,049	0,044	0,122
	(,036)	(,036)		(,038)	(,037)	
Credit card and cvc code (4%)	-0,013	0,019	0,004	-0,014	0,023	0,048
	(,023)	(,028)		(,025)	(,030)	

*Note: Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1. The percentages in parentheses are the percentage of respondents of the first control group that were willing to disclose this item.*

The dependent variables items are ordered from most willing to disclose to less willing to disclose. For all groups more than 96 percent of the respondents were willing to share their e-mail. Whereas only 3 percent were willing to share their credit card number and cvc code. At first

sight, none of the coefficients of interests are significant under the 5 percent level. Based on this the alternative hypothesis that consumers are more willing to share information after the GDPR can be rejected.

For most of the items, the coefficients are similar for the models with- and without controls. For 'Date of Birth' the coefficient of being in the treatment group changes from negative to positive for the model with controls.

The coefficients of e-mail, gender, name, address, and credit card and cvc code are negative and smaller than the beta of the second control group. This implies that the consumers in the treatment group are less willing to give this information in comparison to both control groups. E.g. people in the treatment group are on average 6,4 percent less willing to share their name with company "SB". This is significant under the 10 percent level.

For the item phone number, the coefficient of the treatment group is positive, but smaller than the coefficient of the second control group. This implies that consumers in the treatment group are more willing to share their phone number in comparison to the first control group, but less willing to share in comparison with the second control group. This result yields a contradiction, as both control groups should give similar results.

Religious beliefs and political preferences have a negative relationship with the first control group, but a positive relationship with the second. The beta of the treatment group is negative, but larger than the beta of the second control group. This result also yields a contradiction.

Monthly purchases, income and medical data have a positive relationship with both control groups. The beta of the treatment group is positive and larger than the beta of the second control group. Consumers are after the GDPR more willing to share this information.

The coefficients do not show some kind of relationship between the order in which information is most- and less willing to be shared. E.g. the relationship between e-mail and being in the treatment group is negative, but this holds also for credit card and cvc code.

Overall, none of the coefficients of interest are significant under the five percent level so the null hypothesis cannot be rejected. Furthermore, the items show different willingness to share (e.g. positive and negative). There is not a consensus about the willingness to share of consumers and the importance of the items.

Table 5: Regression estimates of the effect of the treatment on the trust in the compliance of the privacy law by companies

Dependent variable: Trust in companies	(1) OLS	(2) Oprobit	(3) OLS	(4) Oprobit
Treatment group	-0.333	-0.191	-0.194	-0.101
	(0.204)	(0.125)	(0.213)	(0.132)
Control group 2	-0.286	-0.136	-0.292	-0.144
	(0.210)	(0.132)	(0.216)	(0.137)
Controls	No	No	Yes	Yes
(Pseudo) R^2	0.008	0.002	0.094	0.027
N	393	393	393	393

*Note: Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.*

At first sight, none of the coefficients of interest are significant. Based on this, there is no evidence that the trust in the compliancy of the privacy law by companies increases after the GDPR. The alternative hypothesis is rejected.

The beta's are negative for both ordered probit models. This implies that the probability of being in the lowest category 'strongly disagree' increases.

The beta of the treatment group is negative and larger than the second control group for the first regression. This means that consumers in the treatment group have less trust in the compliancy of the privacy law by companies in comparison to both control groups.

The fourth model, the treatment group's coefficient is negative but smaller than the coefficient of the second control group. This implies that the treatment group has a negative relationship with the first control group, but a positive relationship with the second control group. This yields a contradiction.

Overall, there is not enough evidence to reject the null hypothesis. The treatment group has a negative relationship with the first control group. For the second control group, it depends on the model with- or without controls.

Table 6: Regression estimates of the effect of the treatment on the trust in the enforcement of the privacy law by the authorities

Dependent variable: Trust in authority	(5) OLS	(6) Oprobit	(7) OLS	(8) Oprobit
Treatment group	-0.171 (0.178)	-0.136 (0.125)	-0.095 (0.184)	-0.066 (0.1390519)
Control group 2	-0.187 (0.186)	-0.111 (0.134)	-0.255 (0.190)	-0.094 (0.133)
Controls	No	No	Yes	Yes
(Pseudo) R^2	0.003	0.001	0.109	0.039
N	393	393	393	393

*Note: Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.*

None of the coefficients are significant. The null hypothesis of consumers having more trust in the enforcement of the privacy law by the authorities cannot be rejected.

The coefficients of the ordered probit models are negative. This means that the probability of being in the treatment group increases for the lowest category 'strongly disagree'.

The beta's for the OLS estimations are negative for the treatment group in comparison with the first control group, but larger than the coefficients of the second control group. This implies that the trust in the enforcement of the privacy law by the authorities decreases after the GDPR in comparison with the first control group. In comparison with the second control group, the trust increases after the GDPR. This yields a contradiction as both groups should give similar results.

Overall, the coefficients are insignificant thus there is not enough evidence to reject the null hypothesis. There is a negative relationship between the treatment group and first control group. There exists a positive result for the second control group.

Table 7: Regression estimates of the effect of the treatment on the trust in the privacy law

Dependent variable: Trust in privacy law	(9) OLS	(10) Oprobit	(12) OLS	(8) Oprobit
Treatment group	0.211 (0.152)	0.188 (0.131)	0.293 (0.163)*	0.261 (0.141)*
Control group 2	-0.141 (0.156)	-0.111 (0.128)	-0.178 (0.161)	-0.154 (0.134)
Controls	No	No	Yes	Yes
(Pseudo) R^2	0.013	0.004	0.096	0.033
N	393	393	393	393

*Note: Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.*

The coefficients of the treatment group are significant under the 10 percent level for the full model with controls. This is above the rejection level of 5 percent significance thus the null hypothesis cannot be rejected.

The beta's of the treatment group are positive in the ordered probit models. The probability of the highest category "strongly agree" increases for the trust in privacy laws after GDPR.

The beta's of the treatment group are positive and larger than the beta's of the second control group. This implies that there exists a positive relationship for the trust in privacy laws after the implementation of the GDPR in comparison with previous privacy laws.

Overall, the trust in privacy laws increases after the GDPR, but the coefficients are not significant under the 5 percent level. There is not enough evidence to reject the null hypothesis.

For robustness, the hypotheses were also tested with the second control group as base group. The results of the treatment effect did not change and were also insignificant. The results are available on request.

VII. Conclusion

This research into the privacy paradox and the General Data Protection Regulation is an exploratory study. A first step in providing an answer to what extent high educated Dutch consumers change their behavior after the GDPR. Using an experiment, an environment was created which could possible lead to measuring the first implication of a law that has not been enforced yet. The population of the experiment was divided in three groups, namely two control groups and one treatment group. The difference between these groups was that the first control group got information about the current legislation. The second control group did not receive any additional information. The treatment group received information about the new privacy law.

The null hypothesis of willingness to share information is not rejected due to insignificance. The models show different willingness to share for the treatment group. This depended on the item. There was no consensus between the willingness to share and the importance of the item. The theory of the privacy paradox predicted that consumers should be more willing to disclose information after the GDPR. Unfortunately the results did not verify this.

Second hypothesis found ambiguous results. The trust in the compliance of the privacy law by companies increased after the GDPR for the first control group. However, in comparison with the second control group the results were ambiguous. The results were insignificant thus the null hypothesis was not rejected.

The third hypothesis of the trust in the enforcement of the privacy law by authorities gave insignificant results and was rejected. The results also showed ambiguous results. For the first group the trust declined after the GDPR. For the second control group the trust increased after the new privacy law.

The trust in privacy law increased after the GDPR for both control groups. The coefficients of the full model were significant under the 10 percent level, which was not enough to accept the alternative hypothesis. The alternative hypothesis of more trust in privacy law was rejected.

This research tries to provide the first implication of the General Data Protection Regulation and the privacy paradox. At this moment, there is not enough evidence found to conclude that consumers change their behavior after the General Data Protection Regulation.

Limitations

One of the major limitation of this research is that the experiment was held using a questionnaire. The groups only differed in the information provided about privacy law. It is possible that a respondent did not read the information well. This would lead to a relative weak treatment effect. For the respondents it is difficult to fill in a questionnaire as if the GDPR has been enforced, although a conformation question was asked. They do not know the implications of the new law, so it is possible that they fill their answers biased to the current legislation.

As described in the literature review, this research only observes stated preferences, due the lack of financial resources. Stating that you are concerned about your privacy is relatively easy. Acting like it when incentivized is quite different. The real behavior of the respondents was hard to capture in this research.

The survey was spread through social media channels and emailing KPMG colleagues of the author. The population sample is not selected at random. This led to an overestimation of respondents from Surinam. It also led to an underestimation of respondents of Morocco, Turkey and other countries. It is possible that this biased the results. The external validity may be relatively low. It could be that there might be other results with a representative random sample.

It is difficult to capture the behavior of consumers. It could depend on all kinds of things. Maybe the respondents differ in characteristics that were not measured. For example, having a bad experience with a company regarding their privacy. Also, consumers may appreciate the individual characteristics of companies. This makes it difficult to answer the behavior for all companies.

Recommendations

Measuring real behavior instead of stated preferences is not difficult. Further studies could give consumers an incentive. This could lead to different results and there might be signs of a privacy paradox. Consumers might be willing to disclose personal information in exchange for a reward. Previous literature solved the stated preferences issue through this research design.

It is difficult to provide implications of a law that has not been enforced yet. Further studies could wait until 25 May 2018. Then they can really observe the implication of the new privacy law. It is also possible to conduct a before and after design. Before the GDPR has been enforced, it is possible to conduct an experiment with incentives to capture the real behavior before the GDPR. After the GDPR has been enforced, redo the experiment. This might lead to a finding of the privacy paradox.

It would be interesting to conduct this research as a panel study. Conduct an experiment before and, a couple years, after the GDPR has been enforced. This design makes it possible to measure the effect of the GDPR on the behavior of consumers.

The GDPR is a law that holds for every company who is processing personal data of EU citizens. It does not matter where the company is based. It would be interesting to conduct this research for other countries. This might give different results. It is also recommended to have a representative sample of the population.

VIII. Bibliography

Allen, I. E., & Seaman, C. A. (2007). Lkert Scales and Data Analyses. *Quality Progress, 40*(7), 64.

Athey, S., Catalini, C., & Tucker, C. (2017). The Digital Privacy Paradox: Small Money, Small Costs, Small Talk. *Working Paper National Bureau of Ecnomic Research*.

Autoriteit Persoonsgegevens. (2016a, Juli 6). *Boetebeleidsregels Autoriteit Persoonsgegevens.* Retrieved from Autoriteit Persoonsgegevens: https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/beleidsregels_van_de_autoriteit_persoonsgegevens_van_15_december_2015.pdf

Autoriteit Persoonsgegevens. (2016b). *Meldplicht datalekken*. Retrieved from Autoriteit Persoonsgegevens: https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/beveiliging/meldplicht-datalekken

Barlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organzational research: Determining appropriate sample size in survey research. *Information technology, learning, and performance journal, 19*(1), 43.

Barnes, S. B. (2006, September 4). *A Privacy Paradox: Social Networking in the United States*. Retrieved from First Monday: http://firstmonday.org/ojs/index.php/fm/article/viewArticle/1394/1312%2523

Bart, I. Y., Shankar, V., Sultan, F., & Urban, G. L. (2005). Are the Drivers and Role of Online trust the Same for All Web Sites and Consumers? A Large-Scale Exploratory Empirical Study. *Journal of Marketing, 69*(4), 133-152.

Belanger, F., & Crossler, R. E. (2011). Privacy in the digital age: a review of information privacy research in information systems. *MIS Quearterly*, 1017-1042.

Beresford, A. R., Kübler, D., & Preibusch, S. (2012). Unwillingness to pay for privacy: A field experiment. *Economics Letters, 117*(1), 25-27.

Brown, B. (2001). Studying the internet experience. HP Laboratories Technical Report. *Hewlett Packard Laboratory*.

Centraal Bureau voor de Statistiek. (2016, November 21). *Bevolking naar migratieachtergrond*. Retrieved from Centraal Bureau voor de Statistiek: https://www.cbs.nl/nl-nl/achtergrond/2016/47/bevolking-naar-migratieachtergrond

Centraal Bureau voor de Statistiek. (2017, November 21). *Arbeidsdeelname; binding met de arbeidsmarkt*. Retrieved from Centraal Bureau voor de Statistiek: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=82922NED&D1=1&D2=0&D3=0,9-12&D4=4,9,14,19,24,29,34,39,44,49,54,59,64,69&HDR=T,G2,G1&STB=G3&VW=T

Centraal Bureau voor de Statistiek. (2017, November 9). *Bevolking; hoogst behaald onderwijsniveau; geslacht, leeftijd en migratie*. Retrieved from Statline: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=82275NED&D1=0&D2=0&D3=0&D4=0&D5=0-1,7,11&D6=59,64,69,l&HDR=T,G1,G3,G5&STB=G2,G4&VW=T

Centraal Bureau voor de Statistiek. (2017, November 27). *Bevolkingsteller*. Retrieved from Centraal Bureau voor de Statistiek: https://www.cbs.nl/nl-nl/visualisaties/bevolkingsteller

Centraal Bureau voor de Statistiek. (2017, July 27). *Bevolkingsteller*. Retrieved from Centraal Bureau voor de Statistiek: https://www.cbs.nl/nl-nl/visualisaties/bevolkingsteller

Centraal Bureau voor de Statistiek. (sd). *Hoogopgeleid*. Retrieved from Centraal Bureau voor de Statistiek: https://www.cbs.nl/nl-nl/artikelen/nieuws/2017/36/meer-hoogopgeleiden-in-alle-beroepsklassen/hoogopgeleid

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika, 16*(3), 297-334.

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

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

EU GDPR. (2016a). *How dit we get here?* Retrieved from EUGDPR: http://www.eugdpr.org/how-did-we-get-here-.html

EU GDPR. (2016b). *GDPR Key Changes*. Retrieved from EUGDPR: http://www.eugdpr.org/the-regulation.html

European Union. (1995). *Directive 95/46/EC.* Brussels: Official Journal of the European Union.

European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and the Council. In E. Union. Brussels: Official Journal of the European Union. Retrieved from http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf

Finstad, K. (2010). Response interpolation and scale sensitivity: Evidence against 5-point scales. *Journal of Usability Studies, 5*(3), 104-110.

García-Gómez, P., Bago d'Uva, T., & Riumallo-Herl, C. (2017, September 21). *Quantitative Spatial Analysis: Modelling discrete choice data 2.* Retrieved November 27, 2017, from Blackboard EUR: https://bb-app01.ict.eur.nl/bbcswebdav/pid-199390-dt-content-rid-686813_1/courses/FEM11087-17/slides%20QMAE%202017-18%20choice%20lecture%202%282%29.pdf

Gross, R., & Acquisti, A. (2005). Information revelation and privacy in online social networks. *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, 71-80.

Grossklags, J., & Acquisti, A. (2007). When 25 Cents is Too Much: An Experiment on Willingness-to-Sell and Willingness-To-Protect Personal Information. *WEIS*.

Hann, I. H., Hui, K. L., Lee, S. Y., & PNG, I. P. (2007). Overcoming online information privacy concerns: An infomration-processnig theory approach. *Journal of Management Information Systems, 24*(2), 13-42.

Harris Interactive. (2004). Privacy On & Off the internet: What Consumers Want. Tech. rep.,.

Huberman, B. A., Adar, E., & Fine, L. R. (2005). Valuating Privacy. *IEEE Security & Privacy, 3*(5), 22-25.

Hui, K. L., Teo, H. H., & Lee, S. Y. (2007). The value of privacy assurance: an exploratory field experiment. *MIS Quarterly*, 19-33.

Jenkins, G. D., & Taber, T. D. (1977). A Monte Carlo study of factors affecting three indices of composite scale reliability. *Journal of Applied Psychology, 62(4)*, 392.

John, R. (2010). Likert Items and Scales. Survey Question Bank: Methods Fact Sheet 1. *University of Strathclyde*.

Kahneman, D., & Tversky, A. (1984). Choices, Values and Frames. *American Psychologist*, 341.

Krasnova, H., & Veltri, N. F. (2010). Privacy calculus on social networking sites: Explorative evidence from Germany and USA. *System sciences (HICSS)*, 1-10.

Krohn, F., Luo, X., & Hsu, M. K. (2002). Information privacy and online behaviors. *Journal of Internet Commerce*, 55-69.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the Americal statistical Association, 47*(260), 583-621.

Kumaraguru, P., & Cranor, L. F. (2005). Privacy Indexes: a survey of Westin's studies.

Li, H., Sarathy, R., & Xu, H. (2010). Understanding situational online information disclosure as a privacy calculus. *Journal of Computer Information Systems*, 62-71.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 1-55.

McKelvie, S. J. (1978). Graphic rating scales- How many categories? *British Journal of Psychology, 69(2)*, 185-202.

Motiwalla, L. F., Xiaobai, L. B., & Xiaoping, L. (2014). Privacy Paradox: Does Stated Privacy Concerns Translate Into the Valuation of Personal Information.

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. *The Journal of consumer affairs, 41*(1), 100-126.

OECD. (1980, September 23). *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*. Retrieved from OECD: http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm#part2

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta psuchologica, 104*(1), 1-15.

Schoenbachler, D. D., & Gordon, G. L. (2002). Trust and customer willingness to provide information in database-driven relationship marketing. *Journal of Interactive Marketing, 16*(3), 2-16.

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

Taddicken, M. (2013). The 'Privacy Paradox' in the Social Web: The Impact of Privacy Concerns, Individual Characteristics, and the Perceived Social Relevance on Different Forms of Self-Disclosure. *Journal of Computer-Mediated Communication*.

Taddicken, M. (2014). The 'privacy paradox' in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication, 19*(2), 248-273.

Tilburg University. (2017). *Interne consistentie - Cronbach's alpha*. Retrieved from Tilburg University:

https://www.tilburguniversity.edu/nl/studenten/studie/colleges/spsshelpdesk/edesk/cronbach.htm

UCLA Institute for Digital Research and Education. (2017, November 27). *Probit Regression*. Retrieved from UCLA Institute for Digital Research and Education: https://stats.idre.ucla.edu/stata/dae/probit-regression/

Universiteit Leiden. (2017, Oktober 4). *Nederland in kopgroep privacybescherming*. Retrieved from Universiteit Leiden: https://www.universiteitleiden.nl/nieuws/2017/09/nederland-in-kopgroep-privacybescherming

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

Westin, A. (2003). Retrieved from http://www.privacyexchange.org/iss/surveys/surveybibliograpy603.pdf

Westin, A., & Harris, L. (1999). IBM-Harris Multi-National Consumer Privacy Survey Tech. rep., 1999. *Approximately 5,000 adults of the U.S. Britain and Germany*.

Westin, A., & Opinion Research Corporation. (2000). Public Records and the Responsible Use of Information. Tech. rep., 2000. *Conducted for the Center for Social and Legal Research and sponsered by ChoicePoint, Inc.*

Westin, A., Harris, L., & Associates. (1990). Findings from the Survey - Consumers in the Information Age for Equifax Inc. 2,254 adults of the national public. *Equifax Executive*.

Westin, A., Harris, L., & Associates. (1991). Harris-Equifax Consumer Privacy Survey. Tech. rep., 1991. *Conducted for Equifax Inc. 1,255 adults of the U.S. public*.

Westin, A., Harris, L., & Associates. (1994). Equifax-Harris Consumer Privacy Survey. Tech. rep., 1994. *Conducted for Equifax Inc. 1,005 adults of the U.S. public.*

Westin, A., Harris, L., & Associates. (1996). Equifax-Harris Consumer Privacy Survey. Tech. rep., 1996. *Conducted for Equifax Inc. 1,005 adults of the U.S. public*.

Xu, H., Teo, H. H., Tan, B. C., & Agarwal, R. (2009). The role of push-pull technology in privacy calculus: the case of location-based services. *Journal of Management Information Systems*, 135-174.

IX. Appendices

In the appendices there is an overview of the calculations of the minimum population sample and the Cronbach's alpha. There is also the do file used in statistical software "Stata".

Population sample calculation

At first, this research was focusing on the whole Dutch population and before we could spread the survey, we had to calculate the minimal observations of the population sample, to have a representative sample. We used the formula of Barlett et al. (2001) to calculate the population sample. Their formula is:

$$n \geq \frac{N \cdot z^2 \cdot p \cdot q}{z^2 \cdot p \cdot q + (N-1) \cdot F^2}$$

On 27 November 2017 10:41, there are 17,170,910 (N) inhabitants of the Netherlands (Centraal Bureau voor de Statistiek, 2017). We use the confidence interval of 95 percent, which lead to the z-level of 1,96. As we do not know the accuracy level exactly, we use the default level of p = q = 50 percent. The margin of error level is also used at the default level of F = 5 percent. So our population sample should be at least:

$$n \geq \frac{17{,}170{,}910 \cdot 1.96^2 \cdot 0.5 \cdot 0.5}{1.96^2 \cdot 0.5 \cdot 0.5 + (4{,}019{,}000 - 1) \cdot 0{,}05^2} \approx 384$$

After calculating our population sample, we spread our survey through Facebook, LinkedIn, WhatsApp and emailing over a hundred KPMG colleagues. This led to a total response of 489 individuals. When we took a first look at the data, we noticed that education level of the dataset was highly skewed to high educated citizens. Approximately 78 percent of the respondents had finished at least their study at a University of Applied Sciences. That is why we focused our research on high educated Dutch citizen and we estimated the population sample again. According to data from Statistics Netherlands, in the second quarter of 2017 4,019,000 (N) persons of 15 years and older were high educated (Centraal Bureau voor de Statistiek, 2017)

$$n \geq \frac{4{,}019{,}000 \cdot 1.96^2 \cdot 0.5 \cdot 0.5}{1.96^2 \cdot 0.5 \cdot 0.5 + (4{,}019{,}000 - 1) \cdot 0{,}05^2} \approx 384$$

If the total sample (N) exceeds 1,000,000 observation then the minimum observations of the population sample does not change, namely 384 respondents. When we dropped individuals who have obtained a degree below the degree of the University of Applied Sciences, we have a sample of 396 respondents. We deleted three observations who stated that their age was zero years old. So the total observations used in this research is 393.

Cronbach's Alpha

Table 8: calculation of the Cronbach's Alpha

	Control group 1		Treatment group		Control group 2	
Variable	Alpha if item deleted	Alpha after deleting items	Alpha if item deleted	Alpha after deleting items	Alpha if item deleted	Alpha after deleting items
Q9	0.634	0.675	0.620	0.652	0.613	0.715
Q10	0.642	0.691	0.618	0.661	0.593	0.656
Q11	0.645	0.697	0.662	0.700	0.614	0.723
Q12	0.707		0.715		0.670	
Q13	0.642	0.705	0.638	0.682	0.616	0.707
Q14	0.633	0.734	0.633	0.674	0.598	0.737
Q15	0.691		0.666	0.688	0.694	
Q16	0.699		0.681	0.709	0.696	
Cronbach's alpha	0.692	**0.747**	0.686	**0.715**	0.67	**0.753**

The dependent variable questions are measured with the seven-point Likert scale. To test the reliability of the scale Cronbach's Alpha (1951) can be used. The alpha says something about the internal consistency of the scale used. When the alpha is above the 80 percent level, researches considers this as a highly reliable scale. When the alpha is below the 50 percent level, it is considered as a highly unreliable scale. In social sciences, a Cronbach's alpha between 70 and 80 percent is considered as acceptable (Tilburg University, 2017).

The Cronbach's alpha is calculated for the three groups separately. We assume that the groups should differ in their answers, so we have to estimate the alpha separately. If we estimated the alpha for the three groups combined, it could have led to a lower level of the alpha. When we estimated the alpha for the three groups separately, we found that the alpha is below the 70 percent level and could be increased by deleting questions. For the two control groups we deleted question 12, 15 and 16, which led to an alpha if approximately 75 percent. We deleted only question 12 for the treatment group, which led to an alpha of approximately 71 percent. These alpha's are all above 70 percent and are considered as 'acceptable'.

The questions deleted were: 12) I find the fine for not complying with the privacy law sufficient, 15) Sharing my data will help me access better products and services and 16) Sharing my data will make me more vulnerable to identity theft. These questions were not necessary for this research and are merely asked to obtain extra information about the privacy concerns of the respondents.

Do- file stata

```
clear

use "c:\users\sballa\Work Folders\Desktop\data"

*data corrigeren
replace Q17 = "55" in 198
replace Q17 = "40" in 242
replace Q17 = "0" in 261
replace Q17 = "25" in 357
replace Q17 = "40" in 406
destring Q17, generate(Q28)
drop Q17
gen Q17=Q28
drop Q28
drop if Q17==0
drop if Q19<6
recode Q19 (6=0 "HBO")(7=1 "WO Bachelor")(8=2 "WO Master"), pre(R)
drop Q19
rename RQ19 Q19
drop Q22_6_TEXT~s
drop Q20_5_TEXT
drop Q22_6_TEXT
drop Progress Status
recode Q26 (1=0 0)(2=1 1)(3=2 2)(4=3 3)(5=4 "4 of meer"), pre(R)
drop Q26 Duration__in_seconds
rename RQ26 Q26
order Q3 Q5 Q6 Q7 Q8_1 Q8_2 Q8_3 Q8_4 Q8_5 Q8_6 Q8_7 Q8_8 Q8_9 Q8_10 Q8_11 Q8_12 Q9
Q10 Q11 Q12 Q13 Q14 Q15 Q16 Q17 Q18 Q24 Q26 Q19 Q20 Q21 Q22

*leeftijd in groepen
gen Q17_=0
replace Q17_=1 if Q17>=20&Q17<30
replace Q17_=2 if Q17>=30&Q17<40
replace Q17_=3 if Q17>=40&Q17<50
replace Q17_=4 if Q17>=50&Q17<60
replace Q17_=5 if Q17>60

*create groups
gen Q3_= (Q3<2)
gen Q5_= (Q5<2)
gen Q7_= (Q7<2)

*summary tables
order Q3 Q5 Q7 Q8_1 Q8_2 Q8_3 Q8_4 Q8_5 Q8_6 Q8_7 Q8_8 Q8_9 Q8_10 Q8_11 Q8_12 Q9 Q10
Q11 Q12 Q13 Q14 Q15 Q16 Q17 Q18 Q24 Q26 Q19 Q20 Q21 Q22
sum if Q3_==1
sum if Q5_==1
sum if Q7_==1

*test of equality > Kwallis-test
gen Q5__=2 if Q5_==1
replace Q5__=0 if Q5_!=1
```

```
gen Q7__=3 if Q7_==1
replace Q7__=0 if Q7_!=1
gen Q3_1=Q3_+Q5__+Q7__
drop Q5__ Q7__

kwallis Q17, by (Q3_1)
kwallis Q18, by (Q3_1)
kwallis Q24, by (Q3_1)
kwallis Q26, by (Q3_1)
kwallis Q19, by (Q3_1)
kwallis Q20, by (Q3_1)
kwallis Q21, by (Q3_1)
kwallis Q22, by (Q3_1)

kwallis Q17 if Q7_==0, by (Q5_)
kwallis Q18 if Q7_==0, by (Q5_)
kwallis Q24 if Q7_==0, by (Q5_)
kwallis Q26 if Q7_==0, by (Q5_)
kwallis Q19 if Q7_==0, by (Q5_)
kwallis Q20 if Q7_==0, by (Q5_)
kwallis Q21 if Q7_==0, by (Q5_)
kwallis Q22 if Q7_==0, by (Q5_)

kwallis Q17 if Q5_==0, by (Q3_)
kwallis Q18 if Q5_==0, by (Q3_)
kwallis Q24 if Q5_==0, by (Q3_)
kwallis Q26 if Q5_==0, by (Q3_)
kwallis Q19 if Q5_==0, by (Q3_)
kwallis Q20 if Q5_==0, by (Q3_)
kwallis Q21 if Q5_==0, by (Q3_)
kwallis Q22 if Q5_==0, by (Q3_)

kwallis Q17 if Q3_==0, by (Q5_)
kwallis Q18 if Q3_==0, by (Q5_)
kwallis Q24 if Q3_==0, by (Q5_)
kwallis Q26 if Q3_==0, by (Q5_)
kwallis Q19 if Q3_==0, by (Q5_)
kwallis Q20 if Q3_==0, by (Q5_)
kwallis Q21 if Q3_==0, by (Q5_)
kwallis Q22 if Q3_==0, by (Q5_)


*hypotheses willingness to share information
reg Q8_1 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_9 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_11 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_7 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_5 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_4 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_12 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_8 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_10 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_3 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
reg Q8_2 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
```

reg Q8_6 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust

reg Q8_1 Q5_ Q7_, robust
reg Q8_9 Q5_ Q7_, robust
reg Q8_11 Q5_ Q7_, robust
reg Q8_7 Q5_ Q7_, robust
reg Q8_5 Q5_ Q7_, robust
reg Q8_4 Q5_ Q7_, robust
reg Q8_12 Q5_ Q7_, robust
reg Q8_8 Q5_ Q7_, robust
reg Q8_10 Q5_ Q7_, robust
reg Q8_3 Q5_ Q7_, robust
reg Q8_2 Q5_ Q7_, robust
reg Q8_6 Q5_ Q7_, robust

*hypotheses trust in companies
reg Q9 Q5_ Q7_, robust
oprobit Q9 Q5_ Q7_, robust
reg Q9 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
oprobit Q9 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust

*hypotheses trust in Authority
reg Q10 Q5_ Q7_, robust
oprobit Q10 Q5_ Q7_, robust
reg Q10 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
oprobit Q10 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust

*hypotheses trust in legislation
reg Q13 Q5_ Q7_, robust
oprobit Q13 Q5_ Q7_, robust
reg Q13 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust
oprobit Q13 Q5_ Q7_ i.Q17_ Q18 i.Q19 i.Q20 i.Q21 i.Q22 i.Q24 i.Q26, robust