How Much Is Data Privacy Worth? Insights from a Norwegian Survey Experiment

Jan Ove Bolstad Vasstveit

UNIVERSITETET I BERGEN

Master's Thesis

Department of Information Science and Media Studies

Information Science

University of Bergen

May 3, 2023

Abstract

This thesis examines the value of data privacy from the individual's perspective within Norway, a country governed by the EU's General Data Protection Regulation (GDPR). Despite heightened awareness of data breaches and privacy scandals, accurately valuing personal data remains challenging due to several reasons such as personal preferences, biases, and knowledge levels.

Using a quantitative approach, this research uses a survey experiment through the Norwegian Citizen Panel to test four hypotheses related to the endowment effect, the sensitivity of information, and the impact of concern and information treatments on data privacy valuations. The survey explores how much individuals are willing to pay (WTP) for data deletion and how much they are willing to accept (WTA) for selling their personal data. The findings indicate significant differences between self-reported privacy concerns and actual behaviours, consistent with the privacy paradox phenomenon.

The results show a large range of valuations, with some individuals placing a high price on their data privacy while others are willing to trade privacy for convenience and monetary benefits. This variation presents challenges for policymakers and businesses in making universally acceptable privacy policies and underscores the importance of considering individual differences in data privacy valuations.

This thesis contributes to the ongoing discussion on the economics of privacy by providing empirical evidence from the Norwegian context, highlighting the implications for policymaking and business in a digital economy.

Acknowledgements

I would like to extend my gratitude to my supervisor, Dag Elgesem, for his constructive guidance. Also, many thanks to Erik Knudsen and the DEMODIGI group for their valuable input and feedback regarding the quantitative survey and analysis. I am also grateful to DIGSSCORE for granting me the Citizen Scholarship and allowing me to use The Norwegian Citizen Panel infrastructure for my research.

Jan Ove Bolstad Vasstveit

Bergen, 02.06.23

1. Introduction

In the digital era, personal data has switched from mere information to becoming a valuable part of the modern economy, influencing consumer behaviour, shaping governmental policies worldwide, as well as playing a role in geopolitics. In our part of the world, most individuals use the internet daily, leaving footprints and personal information in the hands of businesses, often tech giants. Individuals' personal information is extremely valuable for tech giants, as their business model is based on users' personal information, which can be used to create targeted ads, share data with third parties, and more. Personal data of all kinds are valuable; if this were not the case, we would not have weekly data breaches, the Cambridge Analytical scandal, and tech giants paying billions in fines for violating regulations. After thousands of new articles about these scandals and data breaches, including books and documentaries, one might assume that data privacy is becoming a more integral part of people's lives and that the price individuals set on data privacy is increasing.

Privacy and data privacy are important for several reasons. Privacy is a prerequisite for personal freedom and autonomy. Without privacy, sensitive information such as health information, financial data, or political attitudes can be misused, leading to discrimination, harassment, or other forms of unfair treatment. In addition, privacy is essential for trust between individuals and institutions, including governments, employers, healthcare and commercial enterprises. When people know that their personal data is handled with respect and in line with the law, they are more inclined to share the necessary information. A well-functioning democracy requires that individuals can express themselves freely and without fear of surveillance. Privacy ensures that people can participate in political activities, express their opinions and participate in public debate without worrying that their personal information will be misused. In an age of increasing technology and digitalisation, unrestricted access to personal data can lead to mass surveillance. Finally, privacy recognises that each person has a right to a private life, where they can develop and explore their identity without intrusion or interference from others.

In recent years, the General Data Protection Regulation (GDPR) has become a key law in data privacy. Created by the European Union, the GDPR sets strict rules for how personal data should be handled and protected. These rules greatly impact how businesses and governments manage personal information. The GDPR requires businesses to be clear and responsible about how they use personal data and to get permission from individuals before using their data. This regulation aims to protect personal privacy and give people more control over their own information. As a result, the GDPR has raised the importance of data privacy and affected how personal data is valued. Companies working in the EU must follow these rules to avoid large fines and keep their customers' trust. This thesis examines how the GDPR influences how people see and value their data privacy, helping us understand the changing world of digital privacy rights.

This thesis examines the value of data privacy from an individual's perspective. Valuating data privacy has been tried before by several researchers, but this has been shown to be harder than one might think. Many factors influence how individuals value their privacy, including personal preferences, biases, variance in knowledge, definitions, context, etc. This makes it hard to evaluate data privacy because data privacy does not seem to have small variances in preferences, like the value of an unspectacular coffee mug, where most people value the mug from 0$ to 10$. There are different kinds of individuals regarding privacy, such as privacy fundamentalists, privacy pragmatics, and those who are unconcerned about privacy. Hence, policymakers must regulate the digital economy with data privacy valuations that range from nothing at all to priceless. This does not only include data privacy, but privacy in general, where we have people referencing George Orwell’s 1984 for the slightest privacy overstep by government officials, and at the same time have people who would rather have government officials monitor everything on the internet to increase safety. Similarly, for businesses, the trade-off between convenience and free services versus data privacy and the risks of personal data falling into the wrong hands is a significant consideration. In the realm of consumer behaviour, we might have people who do not use any of the products that collect data to protect their privacy, and at the same time, have people who press ‘accept cookies’ faster than a knife fight in a phone booth. These diverse attitudes present significant challenges for crafting universally acceptable privacy policies.

In the timeframe of writing this thesis, Meta (the owners of Facebook and Instagram) introduces a paying subscription between $9.99 and 12.00$ a month. When individuals purchase a subscription, their personal data will not be utilised to create targeted ads. However, it will not stop Meta from collecting and using user data for other purposes. Looking at how many users that choose to pay for the subscription can give an indicator of how much data privacy is worth for individuals, even though Meta still collects and uses user data. Though, it does not seem like these statistics are available yet. This was a solution forced on Meta by EU regulators.

Data privacy is not regulated the same way across the world, and the EU is known for having the strictest rules of all regulators. This thesis examines the Norwegian population, where Norway falls under the EU General Data Protection Regulation (GDPR). Hopefully, most businesses handling personal information are aware of GDPR. This regulation lays down rules for protecting individuals regarding their data privacy and ensures that data transfer within EU borders does not violate the fundamental right to privacy. Knowing what data privacy is worth for individuals would greatly help policymakers worldwide, and it would even be valuable for tech giants worldwide. Not knowing how people value their data privacy provides a substantial challenge for policymakers in regulating how tech giants and other businesses can use and utilise this data.

Many researchers have tried to answer what privacy is worth to the general public, and the valuation has implications for policymaking, business, and general public welfare. Valuing privacy is challenging for several reasons: It is context-dependent, it varies across the world, and it is hard to define privacy. Research has shown that people say they are concerned about their privacy but act differently (Athey et al., 2018).

2. Theoretical framework

This chapter will describe the theoretical framework underlying the concepts discussed in this thesis, with a primary focus on the privacy paradox and its implications, as well as privacy cynicism.

2.1 Privacy paradox

In its essence, the privacy paradox is the mismatch between individuals stated concern for privacy and their willingness to disclose personal information for relatively small rewards (Kokolakis, 2017). Colnago et al. (2023) define the privacy paradox as the ‘gap between individuals’ claims of caring about privacy and their actual behaviours”. This inconsistency of individuals concerns for privacy and their actual behaviour online has been studied by several researchers. Individuals tend to report a high concern for their privacy, but when researcher monitor their behaviour, it often shows that their behaviour is not consistent with their self-reported concern for privacy, and that individuals give up privacy relatively easy. A lot of people who are concerned about their privacy may relate to this paradox, when accepting cookies and terms just for convenience, even though one might care greatly for one’s own privacy. Dienlin and Trepte (2015) defines the privacy paradox as follows: “People’s concerns toward privacy are unrelated to the privacy behaviours.” There are other aspects of privacy, for example territorial privacy, but all the research discussed in this thesis is related to information privacy paradox, even though the term privacy paradox is used.

As mentioned above, there has been done substantial research on the topic of the privacy paradox, and there is evidence that both supports the existence of the privacy paradox and evidence that challenge the privacy paradox. In most studies, the concept of the privacy paradox is described as the conflict between individuals' attitudes towards privacy and their actual privacy-related behaviours. However, some researchers draw a comparison between privacy concerns and behaviours instead. It's important to understand that privacy concerns and attitudes are related but are different from each other. Privacy concerns are generally broad and may not be tied to a particular context, whereas privacy attitudes specifically assess one's stance towards certain privacy practices. Context is important, as we will see later. Another important differentiation lies between privacy behaviour and privacy intention. Many studies focus on privacy intentions (like the survey experiment in this thesis) rather than actual behaviours, which may overlook an important aspect of the paradox: the tendency that privacy intentions do not always translate into protective behaviour (Kokolakis, 2017).

Early studies that first demonstrated the privacy paradox looked at individuals' privacy preferences and then their shopping behaviour, both online shopping and physical grocery shopping. Brown (2001) asked individuals about their privacy preferences through in-depth interviews, while Spiekermann et al. (2001) used a questionnaire to find the privacy preferences. Both showed that the online shopping behaviour did not match their pre-stated privacy preferences. Acquisti and Grossklags (2005) gathered survey data showing that people's privacy decisions are often influenced by limited information, limited reasoning abilities, and biases. Despite 89% expressing significant privacy concerns, over 21% shared their social security numbers for benefits like discounts, and more than 28% gave out their phone numbers for various reasons. Norberg et al. (2007) conducted a study where they first asked to specify which pieces of personal information, they were willing to disclose. After several weeks, a market researcher asked the participants to disclose the same pieces of personal information. The results showed that students were willing to disclose significantly more information to the market researcher than they previously said they would. The results that show the existence of the privacy paradox have been demonstrated by several others as well. In the context of social network sites, Tufekci (2008) found little to no relationship between self-disclosed privacy concerns, using a questionnaire, and behaviour on social network sites. However, the study did show that students that had high privacy concerns, were less likely to create a profile on a social network site, but when they did create a profile, this correlation diminished. The study also found that some students do try to limit their visibility by using nicknames and restricting profile visibility, but do not moderate their disclosure behaviour. This was also showed by Reynolds et al. (2011) where they found little correlation between participants' concern about privacy on Facebook and their posting behaviour. This is highly relevant since posting behaviour (and general behaviour on social network sites) are some of the things Facebook collect data from (Facebook).

Some evidence challenges the notion of the privacy paradox, or at least moderate the results already mentioned. Studies show that individuals disclose information when there are benefits to do so, but at the same time concerned about how this information is handled. Individuals are especially concern about secondary use of their information, and this concern leads to behaviour caution (Kokolakis, 2017). A study using a questionnaire and not factual purchasing data found that there is a relationship between privacy concern and privacy behaviour using two online surveys to look at the influence of price and secondary use on purchase likelihood. They found that the absence of secondary disclosure (i.e. the sharing of personal information to other marketers or third parties) had stronger influence than a 10% price cut (D'Souza & Phelps, 2009). Tsai et al. (2011) carried out a study in two parts, incorporating both a survey on online privacy concerns and a practical online shopping test. This test featured a shopping search engine that clearly presented information on privacy policies. The findings indicated that when consumers are given good privacy information, they actively consider this in their decision-making, often choosing to buy from online retailers who provide moderate to high privacy safeguards. This led to the conclusion that consumers are prepared to spend more for enhanced privacy protections. As mentioned above, younger individuals tend to use different strategies to restrict their personal information on social network sites, like giving false names and misinformation, adjusting privacy settings, restricting friend requests, and deleting tags and photos (Kokolakis, 2017). This favours the argument that privacy concern leads to cautious privacy behaviour. Dienlin and Trepte (2015) replicate the research done on privacy concerns and privacy behaviour, where many of the studies found that privacy concern was not associated with specific privacy behaviour on social network sites. Dienlin and Trepte (2015) conclude that privacy concerns and behaviour were mostly unrelated on social network sites and that privacy concerns cannot sufficiently predict privacy behaviour on social network sites. According to Dienlin and Trepte (2015), the "privacy paradox can still be found—as long as it is investigated as suggested and outlined earlier." However, using a new approach, Dienlin and Trepte (2015) argues that the privacy paradox is a relic of the past. In this study, they differentiated between informational, social, and psychological privacy attitudes and between the informational, social, and psychological privacy behaviour. They found that when differentiating these dimensions and differentiating between privacy concerns and privacy attitudes, privacy attitudes do, in fact, affect privacy behaviour.

While substantial evidence supports the privacy paradox, some studies suggest that the relationship between privacy concerns and behaviours is more nuanced. Context, awareness, and individual differences play significant roles, indicating that the privacy paradox may not be as clear-cut as initially thought. This mixed picture underscores the complexity of the privacy paradox and suggests that more research is needed to fully understand the factors that influence privacy-related behaviours.

Kokolakis (2017) summarizes five research areas that may interpreter the gap between privacy concerns and privacy behaviour: i) privacy calculus theory, (ii) social theory, (iii) cognitive biases and heuristics in decision-making, (iv) decision-making under bounded rationality and information asymmetry conditions, and (v) quantum theory homomorphism. Privacy calculus theory propose that individuals' privacy behaviour is a result of the risk reward trade-off between expected loss of privacy and the potential gain of disclosure. The rewards may be convenience, the need for entertainment and social validation among other factors. The risk is reduced privacy. Social theory highlights how online social networks integrate into users' social lives, making information disclosure a necessity despite privacy concerns. This may be relatable for a lot of people that for example choose to stay on Facebook, despite privacy concern, because of their vast online network and the way Facebook is used to communicate (an example may be how online event invitations are used to organize events). Cognitive biases and heuristics in decision-making indicates that decision-making in privacy matters is influenced by cognitive biases like optimism bias, overconfidence and affect heuristic. Affect heuristics involve making quick judgments and decisions based on emotional reactions, which leads to overestimating the risk of thing we don't like and underestimating the risk related to things we like. These biases lead to underestimating risks and overvaluing immediate disclosure benefits. Bounded rationality, incomplete information, and information asymmetries refers to how individuals may lack knowledge of information privacy and that most people are unable to calculate the risk-reward trade-off between disclosure and privacy. Information asymmetry refers to how for example big tech companies like Meta (owner of Facebook) are more aware of how individuals' personal information is used, compared to the people using the product. The last explanation, quantum theory homomorphism, proposes that the timing of the disclosure is relevant and that preferences may be altered at the time a decision is actually made (Kokolakis, 2017).

The privacy paradox, even though challenged, has been evident in many studies, and factors like context and definitions (of privacy terms) seems to influence the results. The privacy paradox is relevant to both businesses and policy makers, making it harder for policy makers to introduce laws and easier for business to gain personal data. The paradox may incentives businesses to collect more data, while it may make it more difficult for policy makers to make laws and policies based on individuals self-reported privacy concerns, because individuals' actions do not match their words.

2.2 Privacy Cynicism

Newer research suggests that some of the explanation of the privacy paradox can be contributed to the concept of privacy cynicism. The concept of privacy cynicism, introduced by Hoffmann et al. (2016), refers to "an attitude of uncertainty, powerlessness, and mistrust toward the handling of personal data by digital platforms, rendering privacy protection subjectively futile." Hoffman et al. (2016) suggest that people who are cynical about privacy believe that companies collecting their data are only looking out for their own interests, not the individuals'. These sceptical individuals might not try very hard to protect their privacy because they feel it will not make a difference. They end up taking risks with their privacy, even though they do not trust these companies, because they feel it is impossible to completely hide their online activities. Building on this research, Oojien et al. (2022) found that individuals with higher levels of privacy cynicism are less likely to engage in behaviours that protect their privacy, using a survey and measuring interaction effects. The same authors also found that high levels of privacy cynicism moderate the effects of vulnerability, benefits, response efficacy and response costs. Vulnerability (i.e. how at-risk individuals feel that their personal information could be misused online) increased privacy protection behaviours when privacy cynicism was low, but no relation was found when privacy cynicism was moderate. When privacy cynicism high, the relationship between vulnerability and privacy protection behaviour turned negative. A higher perceived response efficacy (i.e. how effective individuals think taking certain actions will be in protecting their privacy) resulted in more privacy protection behaviour when privacy cynicism was low but had no significant effect when privacy cynicism was moderate and high. For people who have low cynicism, seeing low benefits in sharing their data or perceiving low costs (like effort, time, and money) in protecting their privacy leads them to protect their privacy more. However, for people who are more cynical about privacy, the opposite happens: even when they see low benefits in sharing data or low costs in protecting their privacy, they don't necessarily protect their privacy more. This shows that cynicism can change how people view the pros and cons of protecting their privacy online (Oojien et al. 2022).

Privacy cynicism is not entirely new and shares similarities with concepts like surveillance realism and privacy apathy, which also describe individuals' resignation towards data collection and surveillance practices. Despite concerns, people might still choose to use the internet, accepting surveillance as an unavoidable part of modern digital life, thus exhibiting a form of resigned pragmatism (Ooijen et al., 2022). This phenomenon may contribute to the mismatch between people's privacy concerns and their actual privacy-protecting actions.

This thesis will investigate the relationship between self-reported privacy concern and WTA (Willingness to Accept) and WTP (Willingness to Pay) to shed light on the privacy paradox in a Norwegian context. WTA refers to the minimum amount of compensation an individual requires to give up their personal data, whereas WTP indicates the maximum amount an individual is willing to pay to protect or keep their personal data private. By looking at these two measures, this thesis aims to understand the economic value individuals place on their data privacy. It will also look at two factors influencing the privacy paradox: information asymmetry and the endowment effect, in addition to privacy cynicism. Information asymmetry occurs when one party has more or better information than the other, which can affect individuals' privacy decisions. The endowment effect refers to the phenomenon where people assign higher value to things they own compared to things they do not own, which may influence how they value their personal data. Additionally, this thesis will examine whether different types of data, including sensitive personal data, affect WTA and WTP. Understanding these factors can provide deeper insights into privacy valuation and the privacy paradox.

3. Literature review

The literature review is structured into four subchapters, where subchapter 1 relates to hypothesis 1 and so on. Each subchapter describes some of the current knowledge, as well as knowledge gaps relating to the hypotheses.

3.1 The endowment effect

Standard economic theory assumes that individuals' willingness to pay for an object is roughly the same as their willingness to accept compensation to give it up. The endowment effect refers to the inclination of people who possess an item to value it higher compared to those who do not own the item. The endowment effect is closely related to "loss aversion," a term from prospect theory, which was developed by Daniel Kahneman and Amos Tversky in 1979 (Tversky & Kahneman, 1979). Prospect theory suggests that losses are perceived as more detrimental than gains are perceived as beneficial. The endowment effect takes this a step further by suggesting that the mere possession of an item causes people to value it more highly than they would if they did not own it. One of the earliest experiments that demonstrated the endowment effect was conducted by Thaler (1980), where participants were given a mug and then asked to set a price for selling it or buying another. The selling price was significantly higher than the buying price, which is not explained by conventional economic theory that assumes rational behaviour. The endowment effect has significant implications for many fields of study, including economics, consumer behaviour, marketing, and negotiations (Thaler, 1980). The endowment effect is proven in many studies and is often manifested either through the exchange paradigm or the valuation paradigm (Morewedge & Giblin, 2015).

There are two often-used ways to measure the endowment effect: the exchange paradigm and the value paradigm. The exchange paradigm demonstrates that when people are given ownership of an object, they tend to value it more highly and are reluctant to trade it, even when given the opportunity to exchange it for another object. The exchange paradigm shows that when two groups are given two different objects (one object for each group), and then have the opportunity to trade with each other, most of the participants choose not to trade. Often in these experiments, there is a third group that can choose between the two different objects, to establish the general preferences (before owning an object) for the two objects. Knetsch (1989) shows an experiment where 89% and 90% of the participants did not want to trade away the object they got, while the third group almost equally preferred the two objects. 56% of the third group preferred one object, while 44% preferred the other. These kinds of experiments prove the endowment effect through the exchange paradigm.

The value paradigm explores the endowment effect by comparing individuals' willingness to accept compensation (WTA) for giving up an item they own versus their willingness to pay (WTP) to acquire an item they do not own. Often in these experiments, half of the participants are given an item (like the coffee mug in Kahneman´s experiment). The other half is asked what price they would pay to buy the coffee mug from those participants who were lucky enough to get them. The individuals acting as sellers specify the lowest amount of money they are willing to accept (WTA) in exchange for the item. Those who do not own the item, have the opportunity to buy the item from the experimenter and indicate the highest amount of money they are willing to pay for the mug (WTP). The valuing paradigm has been used to prove the endowment effect for many items. Still, the ratio between WTA and WTP varies, depending on several factors like experimental design, items in question, context and more.

3.1.2 The WTA/WTP ratio

Since Thaler, Kahneman, and Tversky pioneered the field of behavioural economics in the 1980s, there have been a lot of experiments looking into the endowment effect and the ratio between willingness to pay and willingness to accept. There are too many experiments conducted to give attention to each one. Luckily, Horowitz and McConnel (2002) conducted a review of WTA/WTP studies, looking at 45 different WTA/WTP experiments. Tunçel and Hammitt (2014) did the same thing, building on Horowitz and McConnel’s (2002) work. Tunçel and Hammit (2014) included 76 studies, including 45 from Horowitz and McConnel (2002). These two studies show the mean ratio and the median ratio of the WTA/WTP experiments in the 45 studies included. The mean WTP/WTA ratio in most studies (80%) was larger than the median WTP/WTP ratio. Horowitz and McConnel (2002) found an average ratio of 7.17 with a standard error of 0.93, which means that participants value something they possess 7.17 times more than if they do not possess it. The median ratio was 2.60 (Horowitz & McConnel, 2002). Tunçel and Hammit (2014) found a geometric mean of ratio of 3.28 for all goods. A geometric mean is often smaller than a 'normal' mean and is calculated $\sqrt{X1 * X2 * \ldots * Xn}$.

Horowitz and McConnel (2002) conclude with a relevant point for the present analysis: "*The high observed WTA/WTP ratios do not appear to be experimental artifacts. Our claim is based on the findings that hypothetical or non-incentive compatible experiments do not yield statistically significantly higher ratios; that high ratios are exhibited by a broad-based (i.e., non-student) population; and that familiarity with the experiments does not uniformly lead to lower ratios.*" (Horowitz & McConnel, 2002). Tunçel and Hammitt (2014) point out that *"Consistent with HM* (Horowitz & McConnel), *we find no significant difference in the disparity between studies using real and hypothetical transactions (...)"* (Tunçel and Hammitt, 2014). This suggests that hypothetical questions (like the questions in this thesis) should reflect participants' actual behaviour.

3.1.3 Type of good

Good	Mean RATIO	Standard error	Number of experiments
Public or non-market goods	10.41	2.53	46
Health and safety	10.06	2.28	32
Ordinary private goods	2.92	0.30	59
Lotteries	2.10	0.20	25
Timing	1.95	0.17	39
All goods	7.17	0.93	201
Unknown number of subjects	6.71	Not calculated	6

Figure 1: Type of good and mean ratio. Horowitz and McConnel (2002).

Horowitz and McConnel (2002) indicate that the further a good is from an ordinary private good, the higher the WTA/WTP ratio. An ordinary private good is a tangible, personal item that is privately owned and consumed, such as a household product or personal possession. Health and safety, as well as public/non-market goods, have the highest ratios, followed by ordinary private goods and lotteries, with the lowest ratios for goods involving the timing of receipt. The study suggests that actual money has the smallest ratio, with no significant difference between WTA and WTP when the good is money. In the public and non-market goods category, the lowest ratio is for the tasting of a bitter substance, suggesting it is perceived more like a private good. The ratios per type of good are also found in Tuncel and Hammitt (2014), where they find the same patterns ranging from environmental goods to ordinary good and money-like goods.

WTA/WTP ratio by type of good weighted by N_{ik}.

Good	Geometric mean of ratio	Geometric standard error of ratio	Number of experiments/surveys	Number of studies
Environmental	6.23	1.20	32	13
Health and safety	5.09	1.18	41	11
Other public or non-market[a]	3.93	1.06	66	15
Timing of receipt[b]	2.64	1.09	6	1
Ordinary private	1.63	1.05	116	28
Lotteries	1.56	1.06	53	13
Time (leisure or travel)	1.45	1.21	7	3
All goods	3.28	1.04	337	76

Figure 2: WTA/WTP ratio by type of good. Hammit (2014).

3.1.4 Super endowment effect for personal information

Winegar and Sunstein (2019) applied the abovementioned framework to an experiment on data privacy. Winegar and Sunstein (2019) found a so-called super endowment effect when measuring WTA and WTP in privacy and personal data, recording ratios up to 1:18. Sunstein and Winegar (2019) used 'willingness to pay' and 'willingness to accept' to find the value of digital privacy in the context of personal data. Since there was such a big gap between WTA and WTP, it is hard to see how people value privacy from Winegar and Sunstein (2019). The results of this single-stated preference survey (in the general category) were that participants would pay 5 dollars per month (median) to delete their data from companies like Google and Facebook.

Still, they demanded 80 dollars per month (median) to let the same companies get access to their personal data (Winegar and Sunstein, 2019). The super endowment effect is also shown in another social media context, where Sunstein (2020) asked participants, *"Suppose that you had to pay for the use of Facebook. How much would you be willing to pay, at most, per month?"* and *"Suppose that you are being offered money to stop using Facebook. How much would you have to be paid per month, at a minimum, to make it worth your while to stop using Facebook?"* (Sunstein, 2020). In this experiment, Sunstein (2020) divided the participants into two groups that either got the first or the second question. The answers were surprisingly similar to the WTP/WTA ratio of data privacy. Sunstein (2020) then expanded the study to include several other digital platforms. The results consistently show a super endowment effect on all platforms (median is listed first, then mean):

Instagram: WTP, $5, $21.67; WTA, $100, $102.60
LinkedIn: WTP, $8, $25.71; WTA, $99, $97.80
Pinterest: WTP, $5, $20.97; WTA, $100, $102.92
Reddit: WTP, $10, $27.73; WTA, $99, $97.73
Snapchat: WTP, $5, $24.92; WTA, $100, $106.20
Twitter: WTP, $5, $19.94; WTA, $100, $104.18
WhatsApp: WTP, $10, $34.90; WTA, $100, $101.16
YouTube: WTP, $5, $17.27; WTA, $88, $90.78

Winegar and Sunstein (2019) find it challenging to explain the super endowment effect for personal information but offer several possible explanations or factors to explain the high WTP/WTA ratio. Winegar and Sunstein (2019) question whether it is the same explanation for both the data privacy questions and the use of social media questions. They further question if the WTP in the data privacy experiment is high or low, especially since 14% of participants answered a WTP of zero and that participants do not care about data privacy. And, if we assume the WTP is low, Winegar and Sunstein (2019) argue that it might be a manifestation of a protest answer because participants may feel they already have a right to delete data. The same expressiveness may be found in the relatively high WTA number as well, where Winegar and Sunstein (2019) argue that participants are manifesting 'moral outrage', saying 'no amount is high enough' and 'I would never sell my data' signalling that their dignity is not for sale (Winegar and Sunstein, 2019). A separate point made by Winegar and Sunstein (2019) is that opportunity costs are more prominent in the WTP question, which may influence the disparity in WTP and WTA. This is because when participants (in general) are asked about WTP, they think about different uses for the money they may spend, while this is not always the focus when people are asked about WTA. They therefore conclude that '*there is additional reason to doubt whether a very high median, in response to WTA questions, is sufficiently informative about the welfare effects of data privacy.'* (Winegar and Sunstein, 2019).

3.1.5 Causes for the endowment effect

A different theory that may explain parts of the endowment effect is the theory about evolutionary advantage. People who acquire more through trading, are in a better position to support more offspring compared to those who undervalue their possessions. The tendency to overvalue something for evolutionary advantage might also 'spill over' into stated valuations where participants do not gain anything from overvaluing (like in this survey experiment) (Morewedge and Giblin, 2015).

Plott and Zeiler (2005) argue that some of the endowment effects found in experiments are due to strategic misrepresentation. Strategic misrepresentation is when participants of the experiment misrepresent the value of WTP and WTA, thinking they are in a negotiation situation. This may account for the gap we see between WTP and WTA, but evidence suggests that strategic misrepresentation only accounts for parts of the gap, if any at all. This seems highly related to the evolutionary advantage theory.

Morewedge and Giblin (2015) mentions reference prices as a factor in determining the endowment effect for goods. Reference prices can be external values or values retrieved from memory that can influence the value given in a WTA/WTP hypothetical study. Known market prices can for example be a reference price. These reference prices can also be dynamic, based on the context in which the goods are priced. An example of this is the difference between a cold beer at a high-end restaurant, versus a beer in the grocery store. The "Reference Price Theory" suggests that people compare the actual value of an item to well-known or market prices they've seen, can remember or know. This means that a person who values going to a concert at 500$ (WTP for a ticket = 500$) would not pay 500$ for the ticket when the market price is 300$. Meanwhile, someone who does not want to go to the same concert and would value the concert experience at 100$ (WTA for a ticket = 100$) would not sell her ticket for 100$, when the market price is 300$. This means that "Reference Price Theory" inflates the seller's WTA and reduces the buyer's WTP because neither wants to make a bad deal (Weaver and Frederick, 2012). Weaver and Frederick (2012) suggest that the endowment effect is often best interpreted as an aversion to bad deals rather than an aversion to losing possessions. Morewedge and Giblin (2015) also mentions psychological ownership and attributed sampling bias as possible causes. Psychological ownership refers to how people may increase the valuation of an object even though they do not legally own the object. Psychological ownership can occur from just touching or imagining one owns the object. Attributed sampling bias refers to the different context or situations individuals are in when stating a WTP or WTA (buyer or seller), where individuals may remember the good things about something they own, while thinking about all the negatives when considering buying something (Morewedge & Giblin, 2015).

Some researchers (e.g. Pachur & Scheibehenne, 2012) also question loss aversion as the sole contributor to the endowment effect and suggest biased information processing, a cognitive theory, as a relevant part of the explanation. Nayakankuppam and Mishra (2005) showed that sellers tend to emphasise the positive aspects of the item they're trading and downplay its negative features, whereas buyers view it in the opposite light. Johnson et al. (2007) suggests the same saying ‘People construct values by posing a series of queries whose order differs for sellers and choosers. Because of output interference, these queries retrieve different aspects of the object and the medium of exchange, producing different valuations’ (Johnson et al., 2007).

3.2 Sensitivity of personal information

This chapter investigates privacy valuation, sensitive personal information and contextual influences. It examines studies by Huberman et al. (2005), Cvrcek et al. (2006), and Tsai et al.

(2011) to understand how individuals assign value to personal information and how contextual factors shape this valuation.

The GDPR categorises different personal data types and ensures special protection for sensitive data, recognising that not all personal data carries the same level of risk or potential harm if misused. A justification for categorising different types of personal data is that EU regulation categorises data and shapes regulation based on how sensitive the data are through GDPR. EU regulation defines personal data as “(...) any information that relates to an identified or identifiable living individual. Different pieces of information that are collected can lead to identifying a particular person and constitute personal data.” (European Commission, n.d.-a). These data include names, surnames, addresses, location data, email addresses, IP addresses and more. They also differentiate sensitive data from the rest of the personal data. Examples of sensitive data are personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, genetic data, health-related data, and data concerning a person’s sex life or sexual orientation (European Commission, n.d.-b). EU regulators provide extra rules for processing sensitive data and it is generally prohibited to process sensitive data. The regulation makes a clear distinction between general personal data and sensitive personal data, and this categorization is essential because it recognizes the varying levels of risk associated with different types of data and ensures that more stricter protections are in place for more sensitive information. These extra regulations imply that sensitive data should be more valuable to the general population.

It can be hard to differentiate between different kinds of personal data, especially for individuals who do not know what different data types are used for. Also, different types of personal data can be valuated differently from person to person. Sensitive information makes people objectively more vulnerable to misuse of data and discrimination. Also, intuitively, health information is worth more to individuals who have a range of what people would call embarrassing diseases in their journal, compared to people that only have records of a broken arm in their journal. This is as mentioned a factor that regulators must take into consideration when valuing data privacy. Even though different types of data are not worth the same to different people, regulators must try to categorize and differentiate less harmful information from more harmful information. In this chapter, the literature on valuations of different types of data will be reviewed as well as how different types of data are categorized in the regulatory framework.

Some of the literature on the value of privacy also differentiates on different data types. These studies show that valuation changes depending on the personal data/information. Winegar and Sunstein (2019) differentiated the questions they asked participants in four categories of personal information. The first group got a general question with no added information about which data the question referred to. This question made the participants define what 'personal data' implies for themselves. The other categories were demographic, identity and health. In the demographic category, personal data were described as 'name, age, gender, profession, household income, address, picture'. The personal data in the identity category were described as 'age, gender, political affiliation, religion, sexual orientation', while in the health category, the personal data was described as 'age, gender, personality traits, physical and mental health'. Winegar and Sunstein's (2019) experiment showed that the participant's willingness to pay had a median of 5 dollars across all the categories, while the willingness to accept had a median of 50-100 dollars. The median WTA for the 'identity' category was 50 dollars, while the median for the 'health' category was 100 dollars. The other two categories had a median WTA of 77.5 dollars and 80 dollars, respectively. Huberman et al. (2005) are among the foundational studies of how individuals value their personal information. Huberman et al. (2005) showed that the value participants placed on their privacy varied depending on the type and sensitivity of the information. For instance, more sensitive information like personal finances had a higher value than less sensitive information. The studies also showed that sensitive information varies among participants. In this study, weight is an example of how some participants value the same type of information more than others, depending on how favourable the participants perceived their weight. In other words, heavy participants valued their weight information higher compared to those who weighed less (Huberman et al., 2005).

This was also shown by Cvrcek et al. (2006), where participants valued their location data, showing that privacy valuation is context sensitive. The valuations changed based on where the participants were located at the time of the study, for example, at home versus at work (Cvrcek et al., 2006). Tsai et al. (2011) observed that most participants had fewer privacy concerns when purchasing typical online office supplies. However, there was a noticeable reluctance when the products were related to personal values and mental states, such as sexual items or books about depression. The reluctance became particularly noticeable for items suggesting violence, like bullets or books about bomb-making (Tsai et al., 2011).

In summary, the studies by Huberman et al. (2005), Cvrcek et al. (2006), and Tsai et al. (2011) collectively demonstrate that individuals value their personal information differently based on the type, sensitivity, and context of the information. This variation in valuation motivates hypothesis 2, which says that individuals will value their data more if these data are sensitive. For policymakers, confirming that sensitive personal data is more valuable than general personal data, contributes to the knowledge about how regulators should treat different types of personal data.

H2: There is a positive relationship between the sensitivity of information and the price demanded, both WTP and WTA.

3.3 Information treatments and data privacy valuation

This subchapter examines how information treatments affect decision-making regarding data valuation and privacy, aiming to contribute to the literature on this topic and its potential policy implications.

Collis et al. (2023) have investigated how information interventions effect consumer valuations of personal data. Their research, a survey experiment, is conducted on two groups of participants. One of the groups is a YouGov sample aimed to represent the adult American internet user, while the other group is recruited via the Data Dividend Project. The latter is a data advocacy group, where members are expected to be more data- and privacy-conscious. Both groups were asked for the minimum amount of money they would require sharing their entire Facebook data with researchers. After this question, participants were exposed to information treatments. There were two information treatments; one which stated Facebook's revenue projection per North American user for the next three years, and another which stated the monetary compensation some Facebook users received in a settlement following a data misuse incident. All participants were randomly assigned to either the revenue treatment or settlement treatment. Both treatments stated the same amount (400$), just different circumstances. After the treatments, participants were allowed to change their valuations (Collis et al., 2021).

The results from Collis et al. (2021) showed that the majority of participants from the YouGov sample who changed their valuation had WTA valuations below 400$ before the information treatment. For those who already valued their data high, there were no difference in WTA before and after the treatment. Of those who changed their WTA valuations and valued their data lower than 400$ before the treatment, 98.2% changed their valuations to a higher number. The treatments made the distribution of valuations less dispersed, since most participants that changed their valuations increased the amount, while few reduced the amount. The percentage of participants with valuations higher than 400$ increased from 60.9% to 70.1%, while the percentage of participants stating a value of exactly 400$ increased from 1% to 7.3% after the treatments (Collis et al., 2021).

Collis et al. (2021) found that women are more likely to change their valuations in response to the information treatment than men. Black participants are more likely to change their valuations in response to the information treatment than White participants. Using Logistic regression, Collis et al. (2021) found that women revise 28% more than men, while Black participants revise 26% more than white participants. This disparity was also found in low-income vs high-income, where low-income participants were more likely to revise than high-income participants. The information treatments reduced the disparity in distribution of valuations but did not eliminate it (Collis et al., 2021).

The results from the Data Dividend Project sample were similar to the YouGov sample. The probability of revision of valuation in the Data Dividend Project sample were 29.4%, where 58.7% of participants with valuations lower than 400$ changed their valuation after the treatments. All the participants in the Data Dividend Project sample changed their valuations to a higher amount. In both the YouGov sample and the Data Dividend Project sample, participants were more likely to respond to the settlement treatment (12% more likely in both samples)

(Collis et al., 2021). A concern may be that these results are just a manifestation of a standard anchoring effect. Collis et al. (2021) deem this unlikely since an anchoring effect should both increase and decrease valuations toward 400$. The anchoring effect would be expected to reduce the higher valuations, which it did not do. Collis et al. (2021) conclude that the valuations are driven in part both by objective information (treatments) and subjective beliefs about data and data privacy. The latter will be explained in the next chapter.

The main reason for examining information treatments in this survey experiment is to add to the literature about how information affects people's decision making. If information about data use, data collection, and in this case tech giant's revenue, can help people more accurately valuate their privacy this could be integrated in policy to level (or at least make it more level) the playing field between consumers and tech giants. This is especially important since the asymmetrical information gap is bigger between tech giants and marginalized groups, than between not marginalized groups and tech giants, as shown by Collis et al. (2021).

The study by Collis et al. (2023) investigates how different information interventions affect consumer valuations of personal data. The findings indicate that providing specific information about Facebook's revenue projections and monetary compensation following a data misuse incident influenced participants' valuations of their personal data. This study motivates the hypothesis that information treatments can impact individuals' perceptions of the value of their personal data and their privacy concerns. Also, if information treatments can help individuals to value their personal data more accurately, one might reduce the information asymmetry between individuals and tech giants, which has implications for both policymakers and businesses.

H3: Respondents who receive information about how much Facebook earns report a higher WTA than those who do not receive this information.

3.4 Privacy attitudes and privacy valuation

This chapter investigates the topic of privacy attitudes and privacy valuation. It examines international differences in information privacy concerns, and how these concerns are influenced by privacy laws and cultural values.

Bellman et al. (2004) looks at international differences in information privacy concerns. They found that peoples worries about privacy are closely linked to the privacy laws in their country, and that some of the differences in privacy concerns are due to cultural values. Another factor affecting privacy concerns online is how familiar someone is with using the Internet. But as people become more experienced online, this concern tends to decrease (Bellman et al., 2004).

Research conducted by Spiekermann, Grossklags, and Berendt (2001) explored the readiness of 171 individuals to disclose private details when offered purchasing advice and price reductions conversing with a chatbot. Their hypothesis was that privacy concerns would “impede the depth and breadth of truthful online interaction (Spiekermann et al., 2001)”. This hypothesis was not confirmed. They found that “participants displayed a surprising readiness to reveal private and even highly personal information (Spiekermann et al., 2001)”

Rothensee and Spiekermann (2008) investigated how privacy awareness effected participants attitude towards the introduction of RFID (radio frequency identification). Their study shows that privacy awareness is negatively related to emotional appreciation and to acceptance of the services that comes with the new RFID technology. So, the more privacy aware the participants were, the more sceptical they was to accept this new technology. They measured privacy awareness by asking participants to self-disclose their attitudes using answer options the researchers developed in focus group research. The two answer options was i) “To me it is irrelevant if somebody knows what I buy for my daily needs,” and ii) “Generally I want to disclose the least amount of data about myself”. As one can see by the answer options, Rothensee and Spiekermann’s (2008) definition of privacy awareness is almost the same as privacy concern.

Rose (2005) looked at New Zealenders willingness to pay for stronger privacy laws with a higher taxing of the population. Participants got a discrete choice between yes and no, where those who answered yes could state how much they were willing to pay. The results find that the majority (52.5%) said no to the WTP question, i.e. they were not willing to pay any extra taxes for stronger privacy laws, even though most participants stated a high concern for privacy. The mean of concern was 4.2, where 5 was 'very concerned'. For those who declined any tax raises for stronger privacy laws, 37.5% said they did not think stronger data protection laws would decrease misuse of personal data, and approximately 22% said "I don't think the money would be used for data protection". These two explanations for not willing to pay for stricter privacy laws is consistent with the privacy cynicism explanation of the privacy paradox. The results for those who were willing to pay higher taxes for stronger laws showed that the median WTP was approximately 28$. 20% of those who were not willing to pay any taxes for stronger privacy laws said that they did not have enough money (Rose, 2005).

Beresford et al. (2012) found that consumers concern for privacy does not always translate into their purchasing decisions. In a field experiment, they measured the willingness to pay for privacy by offering participants the opportunity to purchase a single DVD from one of two identical online stores, were one of the online stores required participants to reveal more personal information than the other. When the store that requested more information offered DVDs at a one Euro discount, the majority of consumers chose the less expensive option, revealing a preference for cost savings over privacy. When the pricing was the same between the two online stores, participants were equally likely to purchase from either store, suggesting that pricing plays a significant role in consumers privacy related decisions.

The results from Rose (2005) and Beresford (2012) is highly relevant also in the context of the privacy paradox, showing how privacy concern does not always reflect privacy behaviour. Testing the relationship between self-reported privacy concern and hypothetical WTA can help to expand on the literature of the privacy paradox, as well as investigating the differences between the Norwegian population compared to other populations.

The literature looks at privacy attitudes and privacy valuation, examining international differences in information privacy concerns, and how these concerns are influenced by privacy laws and cultural values. Bellman et al. (2004) found that people's worries about privacy are closely linked to the privacy laws in their country and that some differences in privacy concerns are due to cultural values. Additionally, familiarity with using the Internet affects privacy concerns online, with concerns tending to decrease as people become more experienced online. The studies mentioned in this chapter, as well as the litterature about the privacy paradox and privacy cynicism motivates the fourth hypothesis.

H4: There is a positive relationship between concern for privacy and the pricing of personal information, both WTP and WTA.

4. Methodology

The methodology chapter focuses on the analysis of the survey experiment data. It involves applying weights using the 'survey' package and conducting linear regression on various variables in the dataset. The chapter investigates relationships between independent variables such as levels of concern, data sensitivity treatment, and Facebook revenue treatment, and dependent variables WTA and WTP. The analysis uses mostly linear regression to model and measure the strength and direction of these relationships, with a focus on understanding the impact of different treatments on valuations. Additionally, methods for assessing WTA and WTP, such as using median values, are explored in this chapter.

4.1 Research design

The aim of the survey experiment is to test the hypothesis stated in the introduction. To recap, those hypotheses are:

H1: There exists a significant endowment effect
H2: There is a positive relationship between the sensitivity of information and the price demanded, both WTP and WTA.

H3: Respondents who receive information about how much Facebook earns report a higher WTA than those who do not receive this information.
H4: There is a positive relationship between concern for privacy and the pricing of personal information, both WTP and WTA.

As several researchers before me, I have chosen a quantitative approach to my research design. This is suitable for several reasons. For example, all the hypotheses include a quantitative variable (WTA and WTP). Also, the hypotheses aim to test relationships between different variables, which can be tested through statistical analysis of quantitative data. The quantitative approach also gives generalisability when the sample is representative. The quantitative approach is obviously to prefer in this study, without going into further detail about qualitative research design. The research design is also confirmatory and deductive since the study aims to test several hypotheses in a quantitative manner. This thesis uses experiment as a method, which involves manipulating one or more independent variables (treatments) and randomly assigning subjects to different treatments to observe the effects on dependent variables. Using experiments as a research method has several advantages, for example the ability to establish causality through control and manipulation of variables, ensuring high internal validity and accuracy. Experiments enable replication, enhancing the reliability of findings.

4.1.1 Survey experiments

Mutz (2011) defines population-based survey experiments as 'an experiment that is administered to a representative population sample' (Mutz, 2011). A survey experiment grounded in population-based methods aims to gather a group of experimental participants that accurately mirrors the specific target population of interest for a given theory. This could be a nation, a state, an ethnic community, or any other subgroup. The sampled population should reflect the broader group to which the researcher plans to apply the results. Population-based survey experiments are a hybrid research methodology between large-scale observational methods such as survey research, and laboratory based experimental approaches. One of the objects of population-based survey experiments is to address the problem of integrating laboratory experiments to a sample population, which is hard to accomplish if the experiments are done in relatively small labs. Population-based survey methods try to utilize the strengths of both the experiment methodology and the survey methodology, while avoiding the weaknesses (Mutz, 2011).

The best way to measure the value of data is giving people an actual choice to see how much they actually are willing to pay for privacy, or how much they are willing to accept less privacy. Since that approach is extremely expensive and impractical, I would argue that a survey experiment is the next best approach. With a survey experiment, one can ask questions to a representative selection of people to find genuine answers while also doing an experiment without these two approaches being mutually exclusive. For hypothesis #4 (There is a positive relationship between concern for privacy and the pricing of personal information, both WTP and WTA.), there is no experiment, and for the other three hypothesises, the data we get from these question may be used for both testing the three experiment-hypothesises and give value to other aspects, for example the absolute sum in WTP and WTA for Norwegians. The point is that the survey itself has value without the experiments, and even though one include the experiments, the survey still provide valuable insights. Survey experiments are, as with any method, limited in some areas, especially while asking about the valuation of an intangible concept such as privacy, but it is still a very good approach for getting a lot of data out of one survey.

4.2 Testing hypotheses

The analysis of this study will utilize a range of different statistical method to test the different hypotheses. To test if there is a significant endowment effect (H1), other researchers have shown the ratio between WTA and WTP. This is fairly straight forward, where one can for example divide WTA to WTA and achieve the ratio. For it to be an endowment effect, WTA divided by WTP should be a positive number. Others have often used the median to calculate the ratio, which I intend to do as well. The median is the preferred method to analyse the hypothesis but there will also be conducted analysis using the mean. For example, the mean has been shown by Winegar and Sunstein (2019) to be not very useful, mainly because of outliers when participants answer unrealistic high numbers, as a way of saying they would never sell their personal data.

To test H2, there will be used descriptive statistics to summarize the data on sensitivity, WTP, and WTA. This includes calculating means, medians, and standard deviations. Also, visualization can help to understand the data and see with naked eye if there is anything going on. Tests to see if there is correlation will be utilized, depending on the nature of the answers, using Pearson's correlation or Spearman's correlation depending on if the data is normally distributed. Testing H3 will be done using correlation and descriptive statistics between groups that gets treatment and those who do not. It is also interesting to see if the effects of the treatment effect how many respondent answer below the treatment amount. H4 is also testet using correlation and descriptive statistics.

4.3 Validity and reliability

There is two types of validity in experimental research; external validity and internal validity. External validity refers to in which extent the results of the survey experiment can be generalized to a wider population, i.e. how the result of the survey can be applied to the whole population, which in this case is Norwegians over the age of 18. The most common approach to increase external validity is using a representative sample. As mentioned, the Norwegian Citizen Panel aspires to have good representativity, but some groups are underrepresented and some and overrepresented. To account for this, we create weights to increase external validity which leads to better generalizability (de Vaus, 2001). Internal validity refers to the extent to which one can be confident that the observed changes in a dependent variable (i.e. participants valuation of their personal data) are directly caused by the manipulation of the independent variable (i.e. the information treatments about Facebook revenue or types of data collected), and not by other factors. To increase internal validity, the sample was divided into four groups: one without treatments, two groups with one treatment each, and one group with both treatments. Participants were randomly assigned to these groups, which should increase internal validity (de Vaus, 2001). There may be reduced ecological validity because of the hypothetical questions asked in the survey. Participants may find it hard to imagine a situation where one gets paid for one's personal data and may have no idea how much one should get paid. Vice versa for deleting data. But, since none of the big tech companies actually pays someone for their data, it can be though to estimate WTA/WTP regardless, and even though hypothetical question is ideal, they still seem suitable for this specific purpose.

Reliability refers to the consistency and stability of the measurements obtained from the survey experiment. If the survey is reliable, repeating it under the same conditions with the same participants should yield similar results. The measurements (i.e. respondents answer to concern, WTP and WTA) consists of a “true” value and the measurement error. Our goal is to have small measurement errors and values that are close to the “true” value. The closer the answers are to the “true” value the more reliable the experiment (Kimberly & Winterstein, 2008). Steps taken to reduce the measurement errors in this study, revolves mainly around asking understandable and clear questions. This is because the questions and the topic may inherently be hard for some people to wrap their head around, and therefore a risk for less reliable results. To increase reliability, I consulted other researcher who have expertise in crafting survey questions and sought feedback from them regarding understanding and complexity of the question. After consulting other researcher, a pilot test was conducted to test the questions in hand. Feedback from people in all stages of life helped the crafting of questions further. These actions were meant to increase reliability, which I think it did.

4.4 Groups and Questions:

Participants are divided into four different groups, with each group receiving different combinations of treatments before answering two main questions.

Question 1 focuses on how much participants would sell their personal data for if companies like Facebook and Google didn't already have access to it. Question 2 assesses how much participants would pay to have all their personal data deleted from these platforms.

Both questions allow participants to provide a monetary value in NOK (Norwegian Krone) as their response.

Participants are asked about their level of concern regarding companies collecting their personal information.

“How concerned are you that companies like Facebook and Google are collecting personal information about you?”

For this question, I have used a Likert scale, where there are five response options ranging from "Very concerned" to "Not concerned at all." This is a ordinal type of measurement, where there is a direction or ranking of feelings/opinions. Though, one can not assume there is equal intervals between the different answer options. For example, if a participant answer option 2 (“concerned”) on the rank, they are not automatically twice as concerned as those who answer option 4 (“slightly concerned”). The rationale for choosing a 5-point scale is that the five point scale is widely used in social science research and most people are familiar with it. In contrast, a seven-point scale may capture more nuances, but research shows that respondents tend to not select extremes the larger the scale is and there is more confusion regarding the differences between the answer options. Also, a three-point scale may not capture the nuances of the respondent. A five-point scale is balanced approach considering the pros and cons of fewer/more answering options (Jamison, 2024)

Group 1: Receives no treatments.
Group 2: Receives the "Data sensitivity" treatment.
Group 3: Receives the "Facebook revenue" treatment.
Group 4: Receives both the "Facebook revenue" and "Data sensitivity" treatments.

The design enables the researchers to assess the impact of each treatment (alone and in combination) on participants valuation of their personal data. By comparing responses across groups, the study can determine how knowledge of company earnings and understanding the specifics of the data collected influence participants' perceptions and decisions related to their personal data's monetary worth.

In summary, this research design employs a combination of a simple survey and an experiment to measure participants' concerns about data privacy and their valuation of personal data in the context of hypothetical monetary transactions.

4.5 Data collection

The data used in this thesis was collected in the 28th wave of The Norwegian Citizen Panel (NCP). This sub-chapter will describe representativity, survey weights and technical aspects of the data collection. The Norwegian Citizen Panel (NCP) is a research-driven survey on the opinions and attitudes of residents in Norway. NCP is a collaboration between several departments at the Faculty of Social Sciences at the University of Bergen and NORCE. Ideas2evidence is responsible for the panel recruitment, the administration of the panel, and the technical solutions regarding data collection and computing (Skjervheim et al., 2023).

The survey 'Norsk medborgerpanel' was distributed to 26 944 panel members and 10 242 answered the 'Norsk medborgerpanel' in total, which is a response rate of 35.1%. To clarify, 10 242 participants was the total participants for 'Norsk medborgerpanel', while 2017 participants answered the survey used in this thesis. The participants could use a smartphone, tablet or other units capable of running web browsers, where 50% used a smartphone. All participants are members of the Norwegian Citizen Panel. The average time usage for completing the survey was 18.1 minutes. The participants were divided into five groups, where the 2017 in group 5 answered the questions relevant to this thesis. The randomisation procedures are executed live in the questionnaire and randomisation functions are written in JavaScript using mostly Math.random() and Math.floor(). There was a randomisation process both to sort into groups, and to sort the participants of group 5 into four different treatment groups.

4.6 Representativity

The Norwegian Citizen Panel aspires to have perfect representativity of the Norwegian population (for individuals above age 18), but according to Skjervheim et al. (2024), there are two explanations on why some groups may be underrepresented. The first explanation is the "access to and familiarity with the internet (given that a web-based questionnaire was the only response mode made available)". This is related to the age composition of the survey participants, where the older participants become, the probability of having an email address and having the skills to complete an online questionnaire often decrease with age. It is no secret that the older one become, the less digital skills one has. The second explanation is the motivation and interest of the respondents, where motivation and interest increase with the increased level of education, i.e. participants with lower levels of education may be underrepresented. The other variables of information about the participants, in addition to age and education level, is age and geography.

As mentioned, the sampling frame of the survey is the Norwegian population above the age of 18, which leaves us with 4.3 million potential participants. The youngest age group (18-29 years) are underrepresented in the survey where the percentage of population in this age group is 18.8%, but only 4.2% of respondents belongs to this age group. In the youngest group, men are underrepresented compared to women (1.7% vs 2.5%). The age group 30 - 59 years is also somewhat underrepresented, where 44% of respondents belong to this group, while 49.7% of the Norwegian population belongs to this age group. In this group, men are also slightly underrepresented compared to women (21.1% vs 22.9%) The oldest group (60 years and above) is overrepresented in this survey, where the net sample is 51.8%, while only 31.5% of the Norwegian population is aged above 60 years. Men are somewhat overrepresented compared to women in the oldest group (27.8% vs 24%). Skjervheim et al. (2024) mentions that the Norwegian Citizen Panel used to have near perfect representativity with regards to age, but over time the oldest group have become increasingly overrepresented, while the two youngest group have been increasingly underrepresented. Skjerveheim et al. (2024) explains this development by pointing out the membership loyalty, where the younger participants are more likely to stop responding to the Norwegian Citizen Panel after being an active member.

Type	18 - 29 years	30 - 59 years	60 years and above
Population	18.8%	49.7%	31.5%
Net sample	4.2%	44%	51.8%

When it comes to education level, there is a systematic underrepresentation of respondents with little or no education, independent of age and gender. The underrepresentation is especially strong for younger respondents, but are present in all age groups. Individuals whose highest level of education is upper secondary are less represented across all age groups, with the exception of men over 60 years old with this level of education. In contrast, participants with a university or college education are significantly overrepresented among the two oldest age groups, regardless of gender (Skjervheim et al., 2024). In this survey, respondents from Northern Norway are underrepresented, while respondents from Oslo, Eastern Norway and Western Norway are overrepresented. The net sample of Trøndelag and Southern Norway are on level with the population. The geographical levels of underrepresentation and overrepresentation is not as prominent as the age levels, and especially not as prominent as the education levels. 50.6% of the respondents are male, while 49.4% is female (Skjervheim et al., 2024).

Weights were created to adjust for biases by comparing the proportion of each stratum in the population to its proportion in the sample. All the 4 variables mentioned (gender, age, education and geography) was included in creating the weights. The weights values range from minimum of 0.2 to a maximum of 5. The weights are capped between 0.2 and 5 to limit the influence of highly overrepresented/underrepresented strata and reduces the probability of weights being overfitted compared to the demographic composition. One percent of the respondents did not input their education level, and in those cases the weight of the education level was set to 1 (Skjervheim et al., 2024).

4.7 Survey questions

The base for the survey questions is Winegar & Sunstein's (2019) article "How Much is Data Privacy Worth? A Preliminary Investigation". Winegar and Sunstein's (2019) article found a super endowment effect, and this survey experiments aims to compare results from Norwegian citizens with those of American citizens. Hypothesis #1 (There exists a significant endowment effect) and #2 (There is a positive relationship between the sensitivity of information and the price demanded, both WTP and WTA) also aims to compare results with the mentioned article.

Therefore, the starting point for the WTA and WTP questions was a direct translation of Winegar and Sunstein's (2019) article. These questions were then revised in collaboration with DIGSSCORE. The feedback from DIGSSCORE was to simplify the questions and make them easier to understand. Two questions regarding the role of social media in the participants life were removed because they added minimal value to the study. Winegar and Sunstein's (2019) article asks for WTA/WTP per month. This survey experiments do not include 'per month'. The rationale behind this is that I think it is easier for participants to understand how the WTA/WTP question would work in the real world, i.e. the practical arrangements, when asking for a one-time sum. Asking for a one-time sum is also somewhat impractical, mostly because once the data is deleted or once the participants give access, how long would participants give access for that sum, and for how long would it be deleted. And also, after it is deleted, can companies then start to collect personal data again? An alternative would be to ask the willingness to pay for companies NOT to collect the participant's personal data per month. The four hypothesises for this survey experiment are all relative, which means that the absolute sum the participants give is less important than the relationships between the other questions and treatments in the survey experiment. Considering the trade-off between understandable questions and precision, I chose to go for understandable questions. This is to get fewer invalid answers since the exact sum participants answer is less important than how these answers correlate with each other.

4.8 Treatments

The study uses a 2x2 treatment design, which means there are two different pieces of information (treatments) that can be presented to participants in various combinations.

The first treatment is an introduction that informs participants about Facebook's earnings from using personal data for targeted advertising.

The second treatment provides details about the type of data collected, emphasizing age, gender, personality traits, and health information.

The treatments are motivated by hypotheses #2 (There is a positive relationship between the sensitivity of information and the price demanded, both WTP and WTA) and #3 (Respondents who receive information about how much Facebook earns report a higher WTA than those who do not receive this information.). It is hard for people to calculate the value of their privacy since privacy is an intangible concept, and most people have no clue how much companies profit from a user's data. That is why I chose to include a treatment about the actual revenue Facebook make from each user in Europe. I do not think the revenue stated in the treatment will shock participants; it is not especially high or low. It does however give participant a grounding in what realm the reasonable value should be. Of course, participants may value their privacy much higher or lower than Facebook's actual revenue per user even if the actual revenue is given. The rationale behind choosing revenue per year per European user is to get a relatable number for participants. One could choose total revenue for all Facebook operations worldwide and get a enormous number, but then it is not relatable to participants. Also, since most people know that Facebook's revenue is very high, it would add less value to the survey, because it would just be another high, unrelatable number. This thesis aims to examine the real world information effect on users, because it is important to see how correct information about privacy influence behaviour of users. The information about Facebook's revenue is statistics from Statista. Statista provided statistics based on revenue per user per region per quarter. I added the four quarters of 2022 together and converter the amount from EUR to NOK (Norwegian Krone) (Statista, 2024-a).

The treatment for hypothesis #2 is based on Winegar and Sunstein (2019), where they defined personal data in different categories. The number of questions in 'Norsk medborgerpanel' is limited, so I could not include all the categories. This survey experiment only includes the health category from Winegar and Sunstein (2019). There are several reasons for including this treatment. First, a definition of what privacy entails can be helpful for participants and may have the same function as the treatment about revenue, where precise and real information can help participant make a more informed choice. Second, there is a wide definition of what privacy is, and research shows that the value of privacy varies based on what information participants are asked to value or share. Lastly, the valuation of privacy is context-dependent, which means that participants valuation of privacy can vary based on the context participants find themselves in. In the case of health information for example, one might see more disparity between WTA and WTP and also more disparity between participants. An example of context-dependency for health data is that one participant may have a mental disorder and would not like to share that under any circumstances, while another participant may have no health issues and be happy to share that. This is based on common argument against stricter privacy laws, where people say ‘I don’t mind less privacy, I have nothing to hide’ (anecdotal).

4.9 Pilot

Prior to gathering data, the survey underwent pilot testing on both a small and large scale. Throughout its development phase, the survey received thorough testing by ideas2evidence and the team of researchers engaged in the study. The results of the pilot tests were considered adequate and significant technical adjustments were not required. The data collection phase began with a preliminary invitation sent to a selectively chosen group of participants known for their high response rates (a soft launch approach). This was to reduce potential problems in the event of technical issues with the questionnaire. As no such issues were identified or reported, the invitation was extended to the rest of the panel members (Skjervheim et al., 2023).

Even though the survey questions were reduced and simplified, there were still some feedback after the pilot. The pilot gives researchers excellent feedback on a lot of the flawed wording and the potential misunderstandings participants have. The pilot is conducted in DIGSSCORE’s office in Bergen, where some participants came to take the survey while an administrator is talking to them about the questions. In the pilot, participants can voice any misunderstanding, fawlty wording, logical flaws in questions etc., and then the administrator reports back to the researchers with the feedback from participants. The feedback this survey got was mostly based around choice of words and challenging questions to understand.

Participants reacted to the wording ‘your data’ and asked what ‘your data’ entails. ‘Your data’’ was changed to ‘personal information’ and ‘personal data’. The definition of data was a recurring theme throughout all the questions. I chose ‘personal data’ and ‘personal information’ over ‘user data’ after a mini survey among friends and family, where the two former words were more understandable and relatable. Another feedback from the pilot was the answers options for the numerical value answers. Some participants tried to press the arrows, which had intervals of 1 NOK, to state their numerical answers. Then if a participant would want to answer ‘500’, they thought they had to press the arrow 500 times. Therefore, the arrows were removed after the pilot, so the participants only choice was to use the keyboard to type in a numerical value.

There were several problems with this numerical answer mechanism. First, some participants refrained from the question by pressing zero, stating that they would not sell their data at any cost. Second, some participants were unsure of what their number meant, and if they understood the question correctly. Lastly, some participants chose an unlike high number, maybe to refrain from the question, and show that they would never sell their data. The large numbers caused error messages as well. We tried to mitigate these problems by adding a sentence to the answer, so if a participant writes '0', the helping sentence would be 'I would sell my data for 0 kroner (give them for free)'. If participants wrote 500, it would be 'I would sell my data for 500 kroner'. There were also added a limit of 1 000 000 kroner or more, just so as not to have problems with error messages. Answers over 1 million is not very helpful either and would be discarded in the analysis anyway. In conclusion, the answer options look like this (translated from Norwegian to English):

"You can max write inn 15 numbers."

- I would pay ________ kroner to delete my personal data.
- I would sell my personal data for ________ kroner.

The Facebook revenue treatment could have been clearer for some participants. Participants asked questions about where Facebook got their revenue from and if the revenue was exclusively generated by exploiting user data. According to Statista (2004-b), 97% of Facebook's revenue is categorised as marketing/advertisement. Therefore, the treatment was changed from 'Facebook earned 650 kroner per European user in 2022' to 'Facebook earned 650 kroner per European user in 2022 by using personal data for targeted advertising'.

The last change after feedback from the pilot was wording of the first question regarding WTA. Instead of wording like 'what amount would you accept per year to allow these entities access to you personal data', it was changed to 'if these entities did not have access to your personal data, but were willing to pay for your personal data, what would you sell them for?'. The latter may be easier to phatem because every part of the sentence is easy to understand for participants. This change made the survey deviate more from Winegar and Sunstein (2019), because selling your data is different from letting companies access them. This may result in a different WTA than other researchers have found when finding valuating privacy, because the new question is more direct WTA on personal data, not privacy.

4.10 Data analysis

There are several ways to prepare a dataset for analysis. The dataset provided to me by the Norwegian Citizen Panel was already nicely structured for the aim of analysis. Cleaning the data in this case mostly meant deciding whether to remove NULL-values and to what extent. I chose to remove NULL-values independently in each analysis of different hypothesis testing. Meaning, when I looked at WTA and concern, I would only remove NULL-values if one of these two columns had a NULL-value. This was also done in the other scenarios, so there were no removing of all rows that had a NULL-value, and all respondents had answered at at least one question.

Dealing with extreme values is a relevant task in analyzing this dataset of survey results.

To standardise the data for my first hypothesis (H1), I have followed the same thought as Winegar and Sunstein (2019), with some modification. Winegar and Sunstein (2019) had a cut-off point at 25 000$ a month, because this is the monthly income of the 99th percentile in the United States. The average income of the 90th percentile in Norway is 112 460 kr (the 99th percentile could be used in this case as well, but the statistics from Statistics Norway present the data in decile) (Statistisk Sentralbyrå, https://www.ssb.no/statbank/table/12521/tableViewLayout1/). Since the questions in my survey does not state a time frame for payments, i.e. willingness to pay per month or per year, I have used the present value of perpetuity formula, which calculates the present value of for example recurring monthly amounts indefinitely. The formula is Present value = Monthly Income/Discount rate. The discount rate could be sat to 4% annually (the interest I get from having money in the bank, but there are several accepted discount rates), which must be divided by 12 months, since the income is monthly. Present value = 112460 / (0.04/12) = 33 738 337 (round up to 34MNOK). This means that 90% of the population in Norway cannot afford amount above 34MNOK, and the top 10% of income earners has to use all of their income on paying big tech companies, given that respondents are allowed to pay in monthly instalments. To be able to compare WTA and WTP, I will also standardise the WTA column and cap both columns at 34MNOK. I still think this is an absurd high amount, but at least it would be possible to pay for some proportion of the Norwegian population. I will also show how the data look like when all values above 1 million are removed (including 1 million). The amount of 1 million is arbitrary but still it is a cognitive cut off for most, and often when people say 'I would pay a million' it is more a statement that they would pay a lot IF they had the money.

For testing relationships between independent variables and dependent variables, I utilize linear regression using the 'survey' package in R. The 'survey' package is, as the name might suggest, especially suitable for survey analysis. This package is especially useful for survey analysis because it supports weighting, stratification, clustering and more to reduce skewness and to ensure representativity. The 'survey' package enables analysis, including linear regression, using the pre calculated weights in a very simple way. In the analysis of the survey experiment, the 'survey package was mainly used for applying weights and to conduct linear regression on the different variables in the dataset.

Several of the hypothesis investigates relationships between an independent variable (levels of concern, data sensitivity treatment, and Facebook revenue treatment) and a dependent variable (WTA and WTP). To investigate these relationships, I have predominantly used linear regression, which is one of the most fundamental and versatile statistical techniques. Linear regression allows us to model and measure the strength and direction of a relationship between an independent variable and a dependent variable. Using linear regression, we can estimate how much the dependent variable is expected to change when the independent variables change. Linear regression is especially useful in our context, to understand which factors significantly (in a statistical sense) influences how individuals valuate their data privacy. Linear regression allows us to quantify the effect of the treatments on the WTP and WTA. The regression coefficient associated with the treatment variable tells us the average difference in valuation between the treated group and the control group, holding other factors constant. This is particularly useful for understanding the impact of a treatment, and how much valuations is expected to change when introduced with a treatment.

Taking inspiration from previous research, a widespread method to look at WTA and WTP is using the median value. One might also use the mean, but the mean can be too influenced by extreme values, which my dataset has plenty of. The mean is used in some of the cases where it might be relevant.

5. Results

This chapter will present the results of the survey experiment, and describe the data collected. The data gathered offered an extreme variance in WTP and WTA values, especially WTA. A lot of respondents have stated values over 1 billion in the WTA question. There is also a lot of respondents that state a WTP value of 0. This is somewhat true in the WTA question as well.

All but 11 respondents answered the survey question regarding privacy concern. A majority of respondents (excluding those 11 who did not answer) stated that they were 'concerned' or 'very concerned' in the question about companies' collection their personal information. 20% (402 respondents) were 'very concerned', while 33.5% (673 respondents) stated that they were 'concerned'. Meanwhile, only 36 respondents stated that they were 'not concerned at all' (1.8%). 601 respondents (30%) answered 'somewhat concerned', and 295 respondents (14.7%) said they were 'slightly concerned'.

As mentioned, there were huge variance in how the participants responded to the WTA and WTP questions in the survey. 661 (40,4% of those who answered the question) respondents answered 0 in the WTP question, with 267 (17%) answering 0 on the WTA question. 626 (40%) stated over 1 million in the WTA question, and 316 (15.7%) of these said above 1 billion. 41% (671)

answered a WTP from 1 up to 1000, 15.8% from 1001 up to a million, while only 2.7% (44 respondents) answered above a million. The same numbers for WTA are 4.1% (64 respondents) for values from 1 up to and included 1000, and 38.9% (609 respondents) for 1001 and up to 1 million.

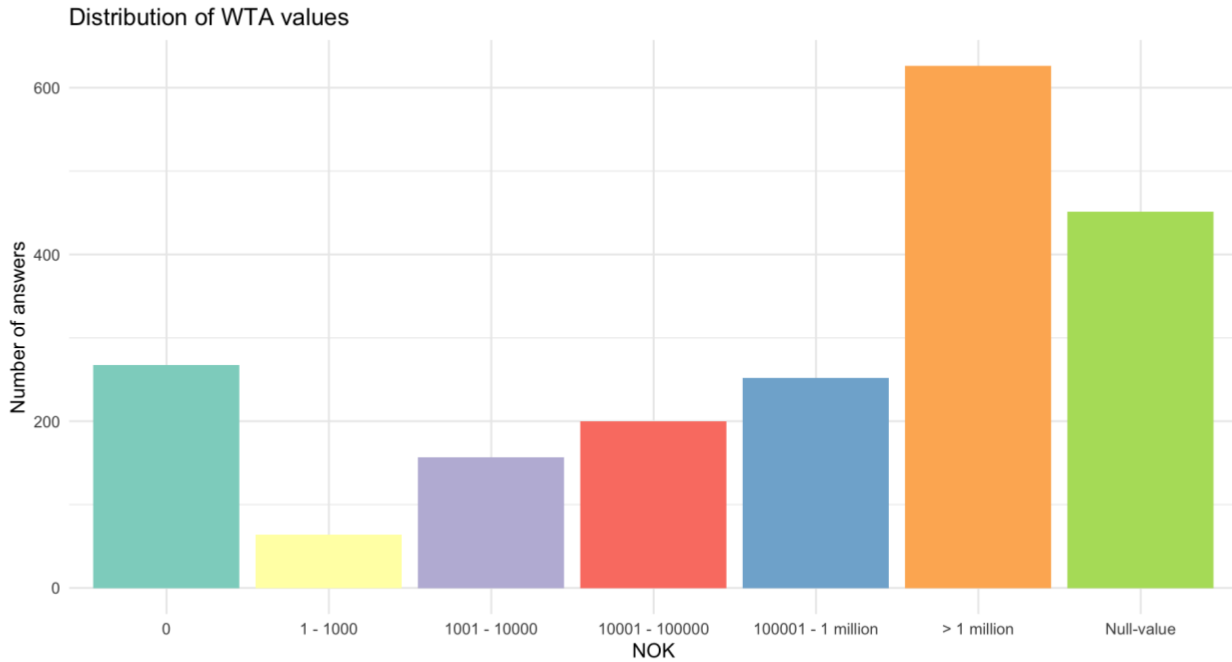

Figure 3: Distribution of WTA values

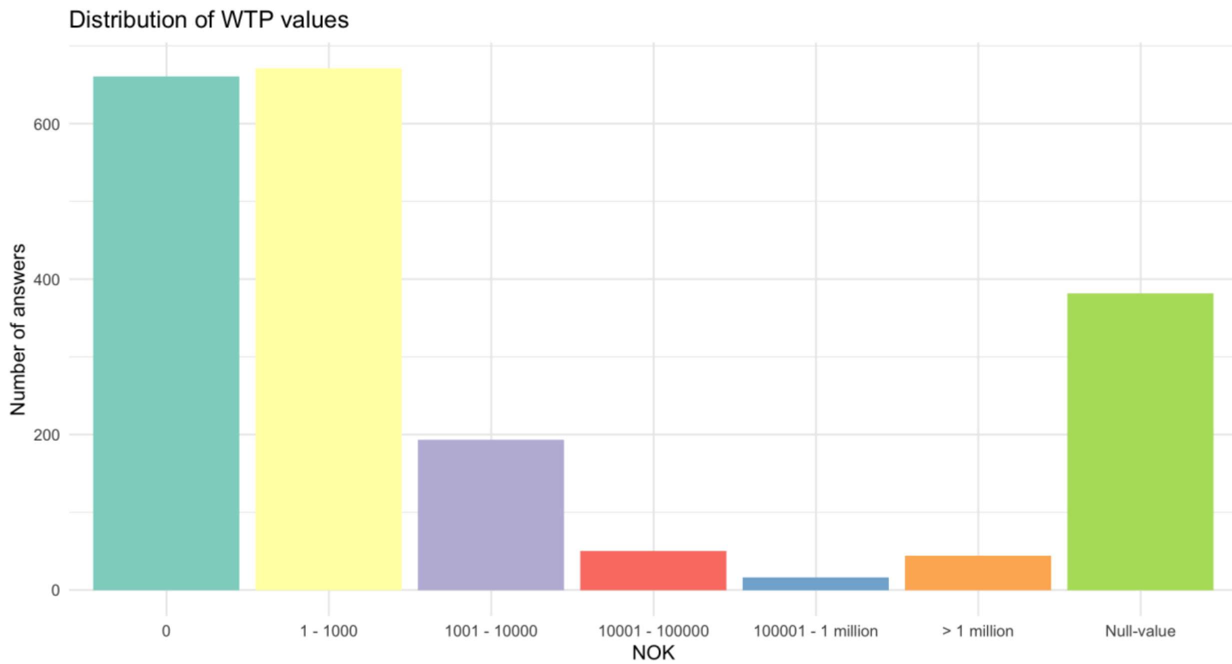

Figure 4: Distribution of WTP values

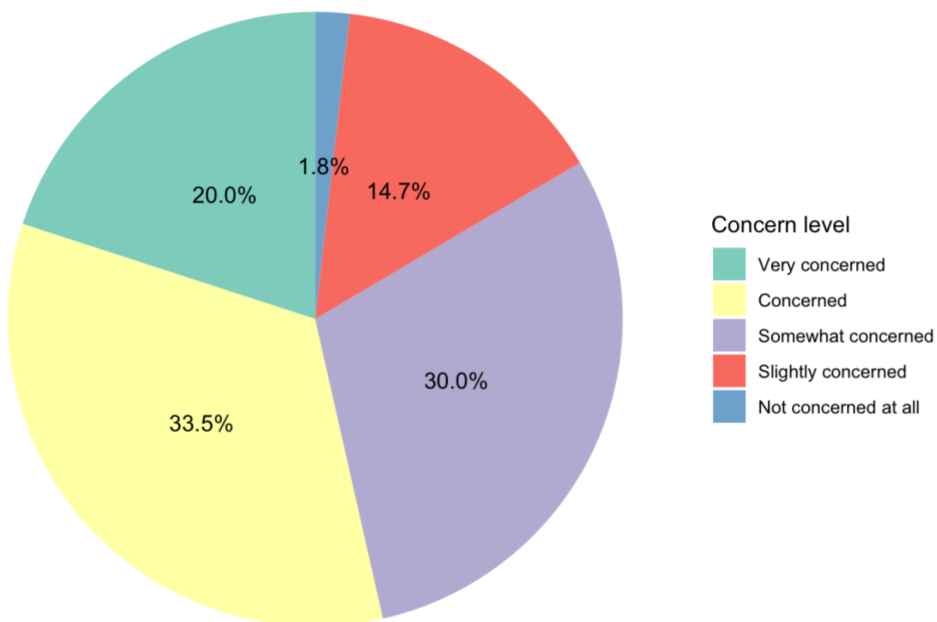

Figure 5: Distribution in % to the question: ''How concerned are you that companies like Facebook and Google are collecting personal information about you?''

5.1 Results for H1

H1: There exists a significant endowment effect

The short answer to H1 is that there exist a significant endowment effect. Also, there exist what Winegar and Sunstein (2019) would call a super endowment effect. Table 1 shows that when including all the answers, the median WTA/WTP ratio is 1000000:100.

Table 1: Summary of responses (unstandardized) in NOK

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
WTA	0	10000	1000000	79000993311532	100000000	999999999999999
WTP	0	0	100	5832415770831	1000	999999999999999

When standardizing the responses to only include values between 0 and 34 million (i.e. when all the answers above 34 million is not included), the median WTA/WTP ratio is reduced to 100000:100 (table 2).

Table 2: Summary of responses (standardized at max = 34 000 000, min = 0) in NOK

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
WTA	0	500	100000	1761261	1000000	25000000
WTP	0	0	100	28912	1000	10000000

Table 3 shows the results when including values between 0 and 1million, where the median WTA/WTP ratio is 9500:100.

Table 3: Summary of responses (standardized at max = 999 999, min = 0) in NOK

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
WTA	0	0	9500	66265	100000	999999
WTP	0	0	100	2175	1000	200000

When calculating the WTA/WTP ratio for each row, the results also show an endowment effect. The median ratio for standardized responses, using this approach, is 200:1, illustrated by the 200 in the median column.

Table 4: Summary of responses when all rows had their own WTA/WTP ratio

Ratio	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Unstandardised	0	10	2000	432004396 26822	2700000	999999999999999
Min = 0 Max =34MNOK	0	1	200	576108	10000	20000000

To find how many individuals that valuated WTA above WTP, I counted the WTA-WTP ratio for each row that had sufficient data (individuals that had both answered the WTP and WTA question). The ratio was made by dividing WTA by WTP, hence ratios above 1 indicate that the respondents would pay less to delete their data than to sell their data. So, the endowment effect is present when the ratio is above 1. The results shows that the majority of ratios are above 1, especially when all the data are included. Removing amounts above 34 million reduced the count of ratios below 1 with 10, the count above 1 with 390, and the count of values of 1 with 10.

Reducing the sample size including only values below 1 million reduced the count above 1 with 556 respondents, but hardly reduces counts of 1 and below.

Table 5: Counts of WTA/WTP ratios below and above 1

Ratio	Count below 1	Count above 1	Count = 1
Unstandardised	100	1194	219
Min = 0 Max =34MNOK	90	804	209
Min = 0 Max = 1MNOK	88	248	203

5.2 Results for H2

H2: There is a positive relationship between the sensitivity of information and the price demanded, both WTP and WTA.

Using linear regression, I found no significant relationships between sensitivity of data and the WTA and WTP values. These results are made comparing values of those who did and did not get the data sensitivity treatment.

Table 6: Linear regression. Standardized (Min = 0 Max=34 000 000) (weighted)

Data sensitivity /WTA	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1469274	453314	3.241	0.00122**

Treatment – Data sensitivity	221854	309366	0.717	0.47344

Table 7: Linear regression. Standardized (Min = 0 Max=34 000 000) (weighted)

Data sensitivity/WTP	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	91578	81943	1.118	0.264
Treatment – Data sensitivity	-6346	53061	-0.120	0.905

Looking at the median and mean, the results shows that the median WTP and WTA value of the standardized repsonses are higher when participants got the data sensitivity treatments. The same applies for the mean.

Table 8: Mean and median with and without treatment. Standardized (min = 0, max=34000000) in NOK

	Median	Mean
WTP (Data sensitivity)	100	82969
WTP (BLANK)	40	78197
WTA (Data sensitivity)	100000	1909746
WTA (BLANK)	50000	1610227

5.3 Results for H3

H3: Respondents who receive information about how much Facebook earns report a higher WTA than those who do not receive this information.

Table 9 and 10 shows the relationship between WTA and Facebook revenue treatment and between WTP and the Facebook revenue treatment. None of them are statistically significant, but the linear regression between WTA and the Facebook revenue treatment is close to the threshold of 0.05. The results of the linear regression indiciates that participants who got the treatment would state a WTA value much lower than those who did not get the treatment (549235 kr less).

Table 9: Linear regression. Standardized (Min = 0 Max=34 000 000) (weighted)

Facebook/WTA	Estimate	Std. Error	t value	P(>|t|)
(Intercept)	2656678	501450	5.298	0.00000014 ***
Treatment Facebook	-549235	307837	-1.784	0.0747

Table 10: Linear regression. Standardized (Min = 0 Max=34 000 000) (weighted)

Facebook/WTA	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2406604	684372	3.517	0.00047 ***
Treatment Facebook	-462111	393296	-1.175	0.24047

Count of values of 650 or under

Excluding all null-values, 149 respondents stated a WTA value of 650kr or below when they did not get the Facebook-treatment, and 153 repsondents stated a WTA value higher than 650kr when getting the Facebook treatment. For WTP, 564 respondents stated a value of 650kr or below when getting the Facebook treatment, while 521 respondets stated a value below 650kr when not getting the Facebook treatment. These numbers are used to discuss the psycological effects of the treatment. 15 respondents answered exactly 650kr on WTP with the treatment, 0 answered 650kr without treatment. 6 respondents answered exactly 650kr on WTA with the treatment, 0 answered 650kr without treatment.

Table 11: Sum of values under 650 NOK with and without Facebook revenue treatment

Sum values under 650kr	**WTA**	**WTP**
NOT Facebook treatment	149	521
Facbook treatment	153	564

15 respondents answered exactly 650kr on WTP with the treatment, 0 answered 650kr without treatment. 6 respondents answered exactly 650kr on WTA with the treatment, 0 answered 650kr without treatment.

The median for WTA when only including respondents who stated values between 1 and 650 was 400 when given Facebook revenue treatment, and 20 when not given Facebook treatment. The mean when given treatment was 314 and 84 when the treatment was not given. For WTP, the median was 100 with and without the treatment, while the mean were 232 given treatment and 160 when not given treatment.

Table 12: WTA and WTP with and without treatment for those who stated a WTA or a WTP between 1 and 650 NOK

1-650kr	Median	Mean

WTP (Facebook)	100	232
WTP (BLANK)	100	160
WTA (Facebook)	400	314
WTA (BLANK)	20	84

4.4 Results for H4

H4: There is a positive relationship between concern for privacy and the pricing of personal information, both WTP and WTA.

The linear regression model does not show a significant relationship, when investigating the link between concern and WTP when including all values under 34MNOK (excluding NULL-values). The table below shows the summary of the linear regression after concerting the different levels of concern into ranked numeric values.

Table 13: Linear regression. Standardized (Min = 0 Max=34 000 000) (weighted)

Concern/WTP	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	96858	52532	1.844	0.0654
Concern	-5794	17602	-0.329	0.7421

Also, the linear regression does not show a statistical significant relationship between concern and WTP for each level of concern with ‘very concerned’ as baseline. None of the different levels of concern are statistically significant different compared to the baseline ‘very concerned’.

Table 14: Linear regression. Standardized (Min = 0 Max=34 000 000) (weighted)

Concern/WTP	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) Very concerned	68616	34739	1.975	0.0484*
Concerned	29279	69161	0.423	0.6721
Somewhat concerned	29692	64588	0.460	0.6458
Slightly concerned	-21981	55535	-0.396	0.6923
Not concerned at all	-10645	65205	-0.163	0.8703

When levels of concern is numerical and ranked, the linear regression shows a strong relationship between WTA and levels of concern. It is estimated that for every level of concerned respondents go down (for example from 'concerned' to ' 'somewhat concerned') they are expected to reduce their WTA value with about half a million NOK.

Table 15: Linear regression. Standardized (Min = 0 Max=34 000 000) (weighted)

Concern/WTA	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3146944	443691	7.093	0.0000000***
Concern	-513543	134861	-3.808	0.000148***

For each of the levels of concern, the linear regression shows that 'somewhat concerned' and 'slightly concerned' has a statistically significant different expected WTA compared to the baseline 'very concerned'. If respondents are 'very concerned', they are expected to state a WTA value of above to million, compared to around a million less if they are ‘slightly concerned’ or ‘somewhat concerned’. Its expected that those who report they are ‘not concerned at all’ is expected to have a WTA of 1396094kr less than those who are ‘very concerned’, but the p-value is just above the standard significance threshold of 0.05.

Table 16: Linear regression. Standardized (Min = 0 Max=34 000 000) (weighted)

Concern/WTA	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept) Very concerned	2279484	408082	5.586	0.0000***
Concerned	254014	550181	0.462	0.64439
Somewhat concerned	-900041	452277	-1.990	0.04683*
Slightly concerned	-1218753	467460	-2.607	0.00925**
Not concerned at all	-1396094	748805	-1.864	0.06252

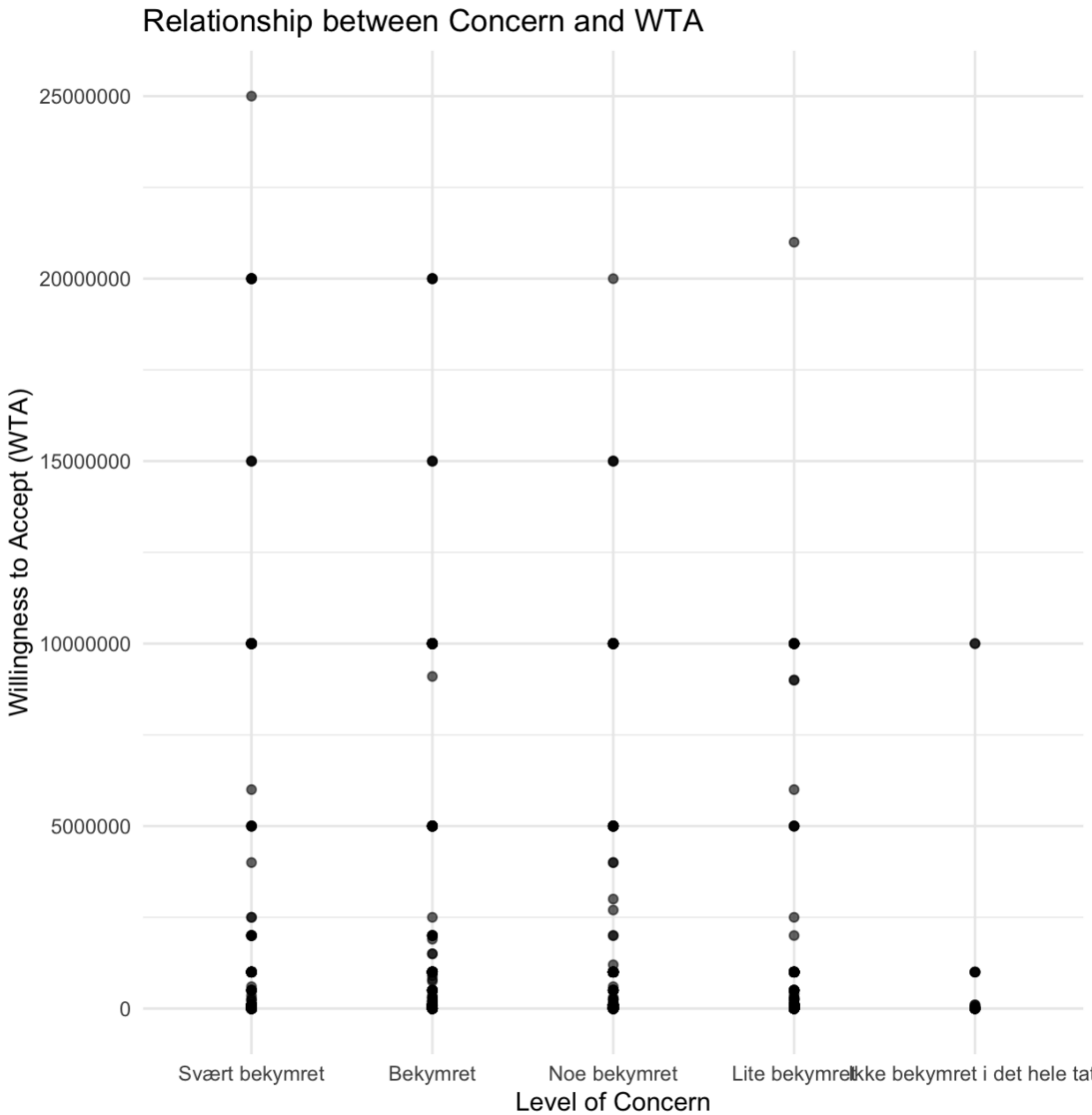

Figure 6: Scatter plot that shows the relationship between concern and WTA. Ranging from very concern on the left to not concerned at all to the right.

6. Discussion

The discussion chapter will examine the findings on privacy concerns and WTA and WTP for data privacy. It will explore the concept of the privacy paradox and how it relates to our research and previous research.

6.1 Hypothesis 1

H1: There exists a significant endowment effect

The first hypothesis was confirming that there exists a significant endowment effect regarding data privacy in the Norwegian population. The endowment effect is clear and present in the survey experiment results; thus we can confirm H1. There are several ways to measure this, and all the methods used in the analysis shows an endowment effect. First, taking the median of the WTA column divided with the median of the WTP column, showed that the endowment effect is present. Second, calculating a WTA/WTP ratio for all rows, i.e. all respondents (that had sufficient data to be included) and then calculation the median for the WTA/WTP ratio also showed an endowment effect. Third, counting WTA/WTP ratios to see how many of the respondents have a ratio of above 1 concluded that the vast majority have a higher WTA than WTP, which is the definition of the endowment effect.

Winegar and Sunstein (2019) called their result ‘a super endowment effect’ when the ratio between WTA and WTP reached 1:18. The results of the analysis of this survey shows an WTA/WTP ratio of 1:1000 using the same method of calculating it. The more the extreme values was included in the calculation, the more extreme the ratio got (1:10000 when all values are present and 1:95 when values over 1MNOK was excluded). The analysis suggests that the super endowment effect also exist in the Norwegian population.

The large WTA/WTP ratios found in the analysis may have several explanations. The most plausible explanation is that these extreme values represent a statement from respondents. This statement may be “I would never sell my data” in the question regarding WTA where values are superficially high. The same goes for the question regarding WTP where the statement might be "I would not pay anything to delete my data, because data privacy is a human right and should be free". These results may indicate that some people perceive personal data as non-market goods. A non-market good is a type of good that is not traded in traditional economic markets, meaning it does not have a defined price or market mechanism where it can be bought or sold. Horowitz and McConnel (2002) found that non-market goods had a higher WTA/WTP ratio than ordinary private goods. The results may indicate that participants treat data privacy as a non-market good, even though personal data is not a non-market good at this time.

These kinds of extreme values may be more frequent in hypothetical scenarios (which this survey experiment is), where there is no actual buying and selling. One can use several WTA values, where one categorizes values above this amount as a statement of 'not for sale'. The analysis used both 34MNOK and 1MNOK. Both these numbers are high and no matter which cut off is chosen, a large proportion answers with a statement. 423 (approx. 27% of the respondents that answered the relevant question) of the respondents stated a higher WTA value than 34MNOK, while 799 (approx. 51%) of the respondent's state 1MNOK or above. Inversely, 661 (approx. 40%) of the respondents stated that they would not pay anything for the deletion of their data (i.e. WTP = 0). The latter number can be interpreted in in two ways. The first interpretation is that they would actually not pay anything to have their data deleted, and that data privacy has no value to them. The second interpretation is that answering zero means that respondents are submitting a statement of something in the like of 'I should not have to pay for my data to be deleted' or that 'deleting my data should be free', as if having data privacy was a human right. To differentiate between these to interpretations, we can look at how many respondents state a WTP of zero AND a WTA of above a million. 289 (approx. 18%) of the respondents stated both zero in the WTP question and above a million in the WTA question. One can assume that these people are in fact stating that 'deleting data should be free', and not thinking that their data privacy is worthless since they would not sell their data for anything less than a million. Some of the high amounts could mean that respondents are thinking of very sensitive data, like social security numbers, when they hear 'personal data', highlighting the importance of GDPR's categorization and protection of such information.

Some of the explanation of the extreme values, especially the large amount of values of zero, may be attributed to privacy cynicism. To recap, privacy cynicism is a feeling of resignation when it comes to data privacy, where people give up on having any data privacy because they feel this is impossible in this day and age. Half of the respondents that stated that they would not pay anything for deleting their data also answered that they are either ‘concerned’ or ‘very concerned’ about companies like Facebook and Google collection their data. Privacy cynicism can also significantly contribute to the high willingness to accept (WTA) numbers observed in this study as well as in other data privacy valuation studies. Privacy cynicism can result in people demanding higher compensation for sharing personal information as a form of protest or moral stance against perceived exploitation by data collectors. Privacy cynics may set high WTA numbers not because they expect to receive that amount, but to signal their strong objection to privacy invasions and that no reasonable monetary compensation can offset the loss of privacy.

Another plausible explanation of why there are so many respondents answering 0 or above a billion is that they did not understand the questions or that they did not take the survey seriously. Since few (2.6%) stated that they would be willing to pay more than 1 million for the deletion of their data, one must assume that most respondents understood the questions and took it seriously.

Loss aversion has been highlighted by several researchers before me as an explanation to the endowment effect. Loss aversion might be a especially important aspect in this survey experiment because of the utility of data. The famous example of Kahneman’s coffee mug for explaining the endowment effect is an example of the utility of the object in hand. A coffee mug do have some utility in which one can drink coffee or other liquids from it. However, in the survey in this thesis, participants are not dealing with a tangible object. Instead, they are considering whether to delete personal data, not purchase it. Even if the question were framed as purchasing their own data, the average person does not get utility from owning their data. Therefore, we must conceptualize this in terms of data privacy: if other entities possess your data, you lack data privacy; if they do not, you maintain it. Thus, the utility of having or owning data privacy is intangible, making loss aversion a factor in how participants value their personal data. So, the utility gain related to purchasing an object (WTP) is less with data privacy; therefore, it may explain low WTP answers compared to WTA answers.

These results might also indicate differences in the population of Norway and the United States of America. As mentioned, Bellmann (2004) found that individuals privacy concerns was closely linked to the privacy laws in different countries. Since Norway fall under the relatively strict privacy regulation of EU, and USA do not, might explain the differences in valuation between this survey experiment and Winegar and Sunstein (2019). Of course, another important factor is the time frame mentioned in the question, where the questions in this survey experiment does not mention a time frame, while Winegar and Sunstein (2019) specify ‘per month’. Another difference between this survey experiment and Winegar and Sunstein (2019), is the wording of the WTA question. Whereas my question asks what participants would sell their data for, Winegar and Sunstein (2019) asks how much participants would demand to let these entities (Facebook, Google, etc.) access their personal data. There might be a significant difference between what people perceive as selling their data compared to let others have access to them. This might increase the number of principled answers like those above a billion, because selling your data feels worse, and in reality, is worse.

6.2 Hypothesis 2

H2: There is a positive relationship between the sensitivity of information and the price demanded, both WTP and WTA.

The results of the linear regression on the relationship between the sensitivity of data and both WTP and WTA showed no statistically significant relationship between treatment groups. The mean and median showed that the median WTP with treatment is 100 and the median WTP without treatment is 40. For WTA, the median with treatment is 100 000 and the median WTA without treatment is 50 000. The mean is greater in both WTA and WTP in groups that were given treatment. Using the same method as Winegar and Sunstein (2019), the results are somewhat consistent. In both their research and in this survey experiment, the median WTA increases when the treatment is introduced. The median WTP of found in Winegar and Sunstein

(2019) did not change when respondents were introduces to treatments, which is the case in my survey experiment.

Looking at the median values, we see that WTA has doubled and WTP has increased by more than double. This suggests that respondents consider the types of data mentioned in the treatment to be of greater value compared to the baseline of just ‘personal data’.

	Median	Mean
WTP (Health data)	100	82969
WTP (BLANK)	40	78197
WTA (Health data)	100000	1909746
WTA (BLANK)	50000	1610227

Another explanation of the lack of relationship between the sensitivity of data and WTA and WTP is that respondents assumes that the sensitive data in the treatment is already included in the question without the treatment. Intuitively, after several documentaries, books and articles about how Facebook use and construct personal data (for example profiling users), one might think that individuals are already aware of which data is used and how Facebook use them. There has been a lot of public education after the Cambridge Analytica scandal in the 2010s. The point is that respondents might already assume the worst about which data Facebook gathers and therefore are unaffected by the treatment.

The results of the linear regression could also indicate that the sensitive data listed in the treatment is not that much more important than just general personal data. This may have implications for regulators as well as businesses because safeguarding specific data type is not something people value more than protection of general data types.

6.3 Hypothesis 3

H3: Respondents who receive information about how much Facebook earns report a higher WTA than those who do not receive this information.

The result of the analysis shows that respondents that gets the information treatment about Facebooks revenue have a tendency to state lower valuations. This is shown in the median of groups that did get the treatment and those who did not. The median of those who did not get the treatment are twice as big as those who did not get the treatment. This was tested comparing group 1 (the group with no treatments) and group 3 (the group with only Facebook treatment). The results stay consistent when including all the groups as well. The estimates from the linear regression showed that those who got the treatment in general stated a WTA of 549235 less than those who did not get the treatment. The results of the linear regression did not meet the conventional threshold of statistical significance, which is a p-value of 0.05. However, with a p-value of 0.07 one might conclude that there is an observable trend, especially with regards to the amount of the estimates and the sample size. My first thought is that these results are just a simple anchoring effect, where the answers from respondent's regress to the number stated in the treatment. The anchoring effect might explain some of it, and even all of it, but there might also be some other factors that can shed light on the results. One aspect is that the treatment informs the respondents about the actual value of their data, making them moderate their answer to more fit the actual value. This might influence respondents to state a bit more reasonable amount, which actually has some root in reality, where they see the Facebook revenue per user as a legitimate benchmark. Also, using the same logic as when standardizing the data, the present value of 650kr a year indefinitely (with a discount rate of 4%) is 16250kr. This can be relevant considering the median went from 100 000 kr without the treatment to 50 000 kr with the treatment, without going into further analysis.

The results of my survey experiment can also be seen in connection and build on the results of Collis et al. (2021). Their results showed that of those who stated WTA valuations of under the reference price, generally chose to adjust their valuations to the reference price or above. The reference price in their study was the results of a settlement where Facebook had to pay 400$ per user, and an expected revenue per user value over 3 years (also 400$). The point is that approximately one third of the participants in Collis et al. (2021) chose to adjust their valuations when presented with informational treatments. The majority of these adjustments was from participants with a lower WTA than stated in the informational treatment. All the respondents that adjusted their valuation and had a previous valuation below the reference price adjusted their values upwards. The results of Collis et al. (2021) could be transferable to my survey experiment results. To see if my results are consistent with Collis et al. (2021), I look at the median and mean of WTA values of 650 or less and above 0, to exclude those who probably not would be affected by the treatment anyways. That leaves us only with 35 respondents, but the results show that the median is 400 for those who got the treatment and 20 for those who did not get the treatment. The mean is 314 with treatment and 84 without. Even though Collis et al. (2021) had the same respondents adjust their valuation (and not two different groups like my survey experiment), the results are still comparable and consistent. The same results are found in the mean of WTP valuations (424 respondents between 1 and 650), but the median is 100 regardless of treatment. A more surprising result that contradicts the results mentioned above is that 153 respondents stated values of 650 or below when given treatment, and 149 stated values 650 or below when not given treatment. Intuition would say that when given the treatment, respondents would state values above the Facebook revenue. This intuition is consistent with the results from Collis et al. (2021), where the proportion of respondents answering above their reference price went from approximately 60% to 70%. For WTP, 564 respondents stated values of 650 or below with the treatment and 521 respondents stated values of 650 or below without the treatment. The WTP results might be easier to intuitively understand, because if respondents get stated the actual market price, they might not automatically feel they should pay above the market price. Quite the opposite, they might feel that they should not pay the market price to delete their data.

6.4 Hypothesis 4

H4: There is a positive relationship between concern for privacy and the pricing of personal information, both WTP and WTA.

The result of the linear regression shows that there is a positive relationship between the concern for privacy and WTA. In general, when the level of concern decreases, respondents state a smaller WTA. The estimate from the linear regression shows a bit more than a half million kroner in reduction for every decreased level of concern. When using 'very concerned' as a baseline, those who responded 'somewhat concerned' and 'slightly concerned' stated lower WTA values of around a million NOK (900041kr and 1218753kr). There is also a trend of those answering 'not concerned at all' having a much lower WTA than those who are 'very concerned' (1396094kr), but this was not statistically significant, though very close to the threshold of 0.05 (p-value = 0.06). The same convincing relationship could not be found with regards to WTP. None of the levels of concern seemed to impact valuation of WTP above the threshold of statistical significance.

The results of the analysis looking at privacy concern and willingness to pay for data privacy is consistent with research on the privacy paradox. The privacy paradox refers to the mismatch between individuals stated concern for privacy and their privacy behaviour. Even though individuals state they are concerned with their privacy, they do not increase their privacy protection behaviour. In the context of this survey experiment, there does not seem to be a connection between how concerned respondents are and how much they would be willing to pay for their privacy, since there was no significant relation between concern and WTP. The findings showing that concern levels influence WTA but not WTP are a perfect illustration of the privacy paradox. Even though people value their data highly enough to demand greater compensation for it (higher WTA), they are not willing to invest similarly in protecting that same data (low WTP).

In a hypothetical scenario where privacy concern and WTP has a perfect positive relationship, one must assume that all individuals have the same income (or a equally wealthy). Equal income is not the case in Norway (or any other country), therefore WTP is most likely influenced by how much respondents earn. This is not intuitively true for privacy concern (it might be, I don't know). Paying 1000kr for deleting data effects poor and rich people very differently, whereas they can all be on the same levels of concern. Individuals can be extremely concerned about privacy, but if they do not have any money to spear then WTP will not reflect concern. This might explain some part of the lack of relationship between concern and WTP. This was also shown by Rose (2005) were 20% of those not willing to pay anything for stricter privacy laws said that they did not have enough money.

7. Conclusion

This thesis set out to examine the valuation of data privacy from an individual's perspective in a Norwegian context. The research used a survey experiment to explore the endowment effect, the sensitivity of information, and the impact of information on data privacy valuations, as well as privacy concerns. The findings reveal significant differences between self-reported privacy concerns and actual behaviours, consistent with the privacy paradox phenomenon.

The results of this study have several important implications for policymakers, businesses, and researchers. Firstly, the significant variation in how individuals value their data privacy underscores the need for personalised and flexible privacy policies. Policymakers should consider these individual differences when making regulations to ensure they are both effective and acceptable to the public. The high willingness to accept (WTA) figures, influenced by privacy cynicism, indicates a mistrust towards organisations handling personal data, suggesting that more transparent and stricter data protection measures could help restore public trust.

For businesses, understanding that consumers place a high intrinsic value on their privacy, even if they do not always act to protect it, highlights the importance of robust data protection practices. Companies that prioritise transparency and user control over personal data may gain a competitive advantage by aligning with consumer values and building trust.

This study opens several areas for further research. Future studies could look closer at the psychological foundations of the endowment effect in the context of intangible goods like data privacy. Exploring the role of cultural and contextual factors in privacy valuations across different regions and demographics would also provide valuable insights. Additionally, studies examining how privacy concerns and behaviours evolve over time, especially in response to new regulations and data breaches, would be beneficial.

Further research could also investigate the effectiveness of different information treatments in changing privacy valuations and behaviours. Looking at the long-term impact of educational and transparency initiatives on consumer attitudes towards data privacy could inform both policy and business practices.

In conclusion, this thesis contributes to the ongoing discussion on the economics of privacy by providing findings from the Norwegian context. It highlights the complex nature of privacy valuations, the significant role of psychological factors like privacy cynicism, and the need for policies and practices that respect and protect individual privacy preferences. As the digital economy continues to evolve, ensuring that privacy remains a protected and valued aspect of our personal and societal well-being will be increasingly important.

References

Acquisti, A., John, L. K., & Loewenstein, G. (2013). What Is Privacy Worth? The Journal of Legal Studies, 42(2), 249–274. https://doi.org/10.1086/671754

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509–514. https://doi.org/10.1126/science.aaa1465

Acquisti, A & Grossklags, J. (2005). Privacy and rationality in individual decision making. IEEE Security & Privacy, 3(1), 26–33. https://doi.org/10.1109/MSP.2005.22

Athey, S., Catalini, C., & Tucker, C. E. (2018, April 8). The digital privacy paradox: Small money, small costs, small talk. MIT Sloan Research Paper No. 5196-17, Stanford University Graduate School of Business Research Paper No. 17-14. https://dx.doi.org/10.2139/ssrn.2916489

Bellman, S., Johnson, E. J., Kobrin, S. J., & Lohse, G. L. (2004). International Differences in Information Privacy Concerns: A Global Survey of Consumers. The Information Society, 20(5), 313–324. https://doi.org/10.1080/01972240490507956

Beresford, A., Kübler, D. F., & Preibusch, S. (2012). Unwillingness to Pay for Privacy: A Field Experiment (IZA Discussion Paper No. 5017). SSRN. https://dx.doi.org/10.2139/ssrn.1634484

Brown, B. (2001). Studying the internet experience. HP laboratories technical report HPL, 49.

Collis, A., Moehring, A., Sen, A., & Acquisti, A. (2021). Information Frictions and Heterogeneity in Valuations of Personal Data (SSRN Scholarly Paper 3974826). https://doi.org/10.2139/ssrn.3974826

Colnago, J., Cranor, L. & Acquisti, A. (2023). Is There a Reverse Privacy Paradox? An Exploratory Analysis of Gaps Between Privacy Perspectives and Privacy-Seeking Behaviors. Proceedings on Privacy Enhancing Technologies. https://doi.org/10.56553/popets-2023-0027

Cvrcek, D., Kumpost, M., Matyas, V., & Danezis, G. (2006). The Value of Location Information (s. 121). https://doi.org/10.1007/978-3-642-04904-0_15

de Vaus, D. (2001). Research design in social research. SAGE Publications Ltd.

Dienlin, T., & Trepte, S. (2015). Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors. European Journal of Social Psychology, 45(3), 285–297. https://doi.org/10.1002/ejsp.2049

D'Souza, G., and Phelps, J. E. (2009), The privacy paradox: The case of secondary disclosure. Review of Marketing Science, 7(1).

European Commission. (n.d.-a). What is personal data? Retrieved April 10, 2024, from https://commission.europa.eu/law/law-topic/data-protection/reform/what-personal-data_en#:~:text=considered%20personal%20data-,Answer,person%2C%20also%20constitute%20personal%20data.

European Commission. (n.d.-b). What personal data is considered sensitive? Retrieved April 10. 2024 from https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en#references

Hoffmann, C. P., Lutz, C., & Ranzini, G. (2016). Privacy cynicism: A new approach to the privacy paradox. Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 10(4), 7. https://doi.org/10.5817/cp2016-4-7

Horowitz, J., & Mcconnell, K. (2000). A review of WTA/WTP studies. Journal of Environmental Economics and Management, 44, 426–447. https://doi.org/10.1006/jeem.2001.1215

Huberman, B. A., Adar, E. & Fine, L. R. (2005). Valuating privacy. IEEE Security & Privacy, 3(5), 22–25. https://doi.org/10.1109/MSP.2005.137

Ivarsflaten, E., Dahlberg, S., Storelv, S., Løvseth, E., Auerbach, K., Bjånesøy, L., Bye, H., Broderstad, T., Böhm, G., Gregersen, T., Knudsen, E., Nordø, Å., Schakel, A., & Tvinnereim, E. (2023) Norwegian Citizen Panel, wave 28 (October - November 2023) [Dataset], v102. Data available from DIGSSCORE, UiB.

Jamieson, S. (2024). Likert scale. In Encyclopædia Britannica. https://www.britannica.com/topic/Likert-Scale

Johnson, E., Häubl, G., & Keinan, A. (2007). Aspects of Endowment: A Query Theory of Value Construction. Journal of experimental psychology. Learning, memory, and cognition, 33, 461–474. https://doi.org/10.1037/0278-7393.33.3.461

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(2), 263–291. https://doi.org/10.2307/1914185

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists, 65(23), 2276–2284. https://doi.org/10.2146/ajhp070364

Knetsch, J. (1989). The Endowment Effect and Evidence of Non-Reversible Indifference Curves. American Economic Review, 79, 1277–1284.

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. Computers & Security, 64, 122–134. https://doi.org/10.1016/j.cose.2015.07.002

Morewedge, C. K., & Giblin, C. E. (2015). Explanations of the endowment effect: An integrative review. Trends in Cognitive Sciences, 19(6), 339–348. https://doi.org/10.1016/j.tics.2015.04.004

Mutz, D. C. (2011). Population-based survey experiments. Princeton University Press.

Nayakankuppam, D., & Mishra, H. (2005). The endowment effect: Rose-tinted and dark-tinted glasses. Journal of Consumer Research, 32(3), 390–395. https://doi.org/10.1086/497550

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. Journal of Consumer Affairs, 41(1), 100–126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

van Ooijen, I., Segijn, C. M., & Opree, S. J. (2024). Privacy Cynicism and its Role in Privacy Decision-Making. Communication Research, 51(2), 146-177. https://doi.org/10.1177/00936502211060984

Pachur, T., & Scheibehenne, B. (2012). Constructing preference from experience: The endowment effect reflected in external information search. Journal of Experimental Psychology: Learning, Memory, and Cognition, 38(4), 1108–1116. https://doi.org/10.1037/a0027637

Plott, C. R., & Zeiler, K. (2005). The Willingness to Pay-Willingness to Accept Gap, the «Endowment Effect,» Subject Misconceptions, and Experimental Procedures for Eliciting Valuations. American Economic Review, 95(3), 530–545. https://doi.org/10.1257/0002828054201387

Prince, J., & Wallsten, S. (2020, January 30). How much is privacy worth around the world and across platforms? TPRC48: The 48th Research Conference on Communication, Information and Internet Policy. https://dx.doi.org/10.2139/ssrn.3528386

Reynolds, B., Venkatanathan, J., Goncalves, J., & Kostakos, V. (2011). Sharing Ephemeral Information in Online Social Networks: Privacy Perceptions and Behaviours (s. 215). https://doi.org/10.1007/978-3-642-23765-2_14

Rose, E. (2005) "Data Users versus Data Subjects: Are Consumers Willing to Pay for Property Rights to Personal Information?," Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 2005, pp. 180c-180c, doi: 10.1109/HICSS.2005.184.

Rothensee, M., & Spiekermann, S. (2008). Between extreme rejection and cautious acceptance: Consumers' reactions to RFID-based IS in retail. Social Science Computer Review, 26(1), 75–86. https://doi.org/10.1177/0894439307307687

Savage, S., & Waldman, D. M. (2013, October 16). The value of online privacy. https://dx.doi.org/10.2139/ssrn.2341311

Skjervheim, Ø., Bjørnebekk, O., Wettergreen, J., Grendal, O., & Stokke, G. (2023). Norwegian Citizen Panel methodology report, wave 28 [Produced by Ideas2evidence].

Spiekermann, S., Grossklags, J., & Berendt, B. (2001). E-privacy in 2nd generation E-commerce: Privacy preferences versus actual behavior. Proceedings of the 3rd ACM conference on Electronic Commerce, 38–47. https://doi.org/10.1145/501158.501163

Statista. (2024-a). Facebook average revenue per user (ARPU) as of 4th quarter 2023, by region. https://www.statista.com/statistics/251328/facebooks-average-revenue-user-by-region/

Statista. (2024-b). Annual revenue generated by Meta Platforms from 2009 to 2023, by segment. https://www.statista.com/statistics/267031/facebooks-annual-revenue-by-segment/

Statistics Norway. (n.d.). 12521: Måneds-lønn, etter desil, statistikkvariabel og år [Monthly salary, by decile, statistical variable and year]. Retrieved Jan 23 2024, from https://www.ssb.no/statbank/table/12521/tableViewLayout1/

Sunstein, C. R. (2020). Valuing Facebook. Behavioural Public Policy, 4(3), 370–381. Cambridge Core. https://doi.org/10.1017/bpp.2018.34

Sunstein, Cass R., How Much Would You Pay to Use Facebook? A Behavioral Perspective (May 21, 2018). Harvard Public Law Working Paper No. 18-30, Available at SSRN: https://ssrn.com/abstract=3173687 or http://dx.doi.org/10.2139/ssrn.3173687

Thaler, R. (1980). Toward a positive theory of consumer choice. Journal of Economic Behavior & Organization, 1(1), 39–60. https://doi.org/10.1016/0167-2681(80)90051-7

Tsai, J. Y., Egelman, S., Cranor, L. and Acquisti, A. (2011), The effect of online privacy information on purchasing behavior: An experimental study. Information Systems Research, 22(2), 254-268.

Tufekci, Z. (2008). Can You See Me Now? Audience and Disclosure Regulation in Online Social Network Sites. Bulletin of Science, Technology & Society, 28. https://doi.org/10.1177/0270467607311484

Tunçel, T., & Hammitt, J. K. (2014). A new meta-analysis on the WTP/WTA disparity. Journal of Environmental Economics and Management, 68(1), 175–187. https://doi.org/10.1016/j.jeem.2014.06.001

Weaver, R., & Frederick, S. (2012). A reference price theory of the endowment effect. Journal of Marketing Research, 49(5), 696–707. https://doi.org/10.1509/jmr.09.0103

Winegar, A. G. & Sunstein, C. R. (2019). How Much Is Data Privacy Worth? A Preliminary Investigation. Journal of Consumer Policy, 42(3), 425–440. https://doi.org/10.1007/s10603-019-09419-y

Appendix

###

SURVEY QUESTIONS

Initial question:
How concerned are you that companies like Facebook and Google are collecting personal information about you?
Answer options:

1. Very concerned
2. Concerned
3. Somewhat concerned
4. Not very concerned
5. Not concerned at all

Treatments (2x2):
● **Facebook revenue:** For example, Facebook earned 650 NOK per European user in 2022 by using personal data for targeted advertising.
Data sensitivity: (including information about your age, gender, personality traits, as well as physical and mental health)

Group 1:

Question 1:

Many apps and websites (Facebook, Instagram, Google, etc.) collect user's personal data. [treatment: no intro] If these entities did not have access to your personal data but offered you money for your personal data, how much money would you sell them for? [treatment: blank]

Answer options: Enter numerical value in NOK

Question 2:

Many apps and websites (such as Facebook, Instagram, and Google) collect the user's personal data. Suppose they would delete the data for a fee. How much would you be willing to pay to delete all your personal data from these places? [treatment: blank]

Answer options: Enter numerical value in NOK

Group 2:

Question 1:

Many apps and websites (such as Facebook, Instagram, and Google) collect the user's personal data. [treatment: no intro] If these entities did not have access to your personal data but offered you money for your personal data, how much money would you sell them for? [treatment: (including information about your age, gender, personality traits, as well as physical and mental health)]

Answer options: Enter numerical value in NOK

Question 2:

Many apps and websites (such as Facebook, Instagram, and Google) collect the user's personal data. Suppose they would delete the data for a fee. How much would you be willing to pay to delete all your personal data from these places? [treatment: (including information about your age, gender, personality traits, as well as physical and mental health)]

Answer options: Enter numerical value in NOK

Group 3:

Question 1:

Many apps and websites (such as Facebook, Instagram, and Google) collect the user's personal data. [treatment: For example, Facebook earned 650 NOK per European user in 2022 by using personal data for targeted advertising.] If these entities did not have access to your personal data but offered you money for your personal data, how much money would you sell them for? [treatment: blank]

Answer options: Enter numerical value in NOK

Question 2:

Many apps and websites (Facebook, Instagram, Google, etc.) collect the user's personal data. Suppose they would delete the data for a fee. How much would you be willing to pay to delete all your personal data from these places? [treatment: blank]

Answer options: Enter numerical value in NOK

Group 4:

Question 1:

Many apps and websites (Facebook, Instagram, Google, etc.) collect the user's personal data. [treatment: For example, Facebook earned 650 NOK per European user in 2022 by using personal data for targeted advertising.] If these entities did not have access to your personal data but offered you money for your personal data, how much money would you sell them for? [treatment: (including information about your age, gender, personality traits, as well as physical and mental health)]

Answer options: Enter numerical value in NOK

Question 2:

Many apps and websites (Facebook, Instagram, Google, etc.) collect the user's personal data. Suppose they would delete the data for a fee. How much would you be willing to pay to delete all your personal data from these places? [treatment: (including information about your age, gender, personality traits, as well as physical and mental health)]

Answer options: Enter numerical value in NOK

##