BI Norwegian Business School - campus Oslo

GRA 19703

Master Thesis

Thesis Master of Science

What are the consequences of externalities and ignorance in online platforms? An economic investigation of current practices and potential

remedies.

Navn: Lars Eric Tveit Ødegård, Elise Mørkeseth

Start: 15.01.2021 09.00

Finish: 01.07.2021 12.00

BI Norwegian Business School
Oslo, Spring 2021

Master Thesis

What are the consequences of externalities and ignorance in online platforms? An economic investigation of current practices and potential remedies.

Lars Eric Tveit Ødegård

Elise Mørkeseth

Supervisor: Gisle James Natvik

MSc Business

Major: Economics

BI NORWEGIAN BUSINESS SCHOOL

This thesis was written as a part of the Master of Science with major in Economics at BI. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Acknowledgements

We would like to extend a special appreciation to Professor Gisle J. Natvik for constructive feedback throughout the process. We would also like to thank the MSc in business with major in Economics class of 2021 for creating a friendly learning environment and outstanding camaraderie.

BI Norwegian Business School

Oslo, June 2021

Lars Eric Tveit Ødegård

Elise Mørkeseth

Abstract

The pace at which online platforms have become an integral part of our everyday life is nothing short of impressive. Many of such platforms have experienced high growth in recent years which highlights the economic value of data. Using a simple economic model, we investigate the consequences of negative externalities in the market for online services and their welfare implications. We find that the consumers do not internalise the social cost of consumption, and hence the social planner suggests a lower level of consumption than in the decentralised market. Consequently, we investigate whether Pigouvian taxation can replicate the planner solution. Despite this, there seems to be a dichotomy between people's privacy attitudes and their behaviour, which raises questions regarding the efficiency in the market for such online services. Therefore, we investigate whether ignorance among consumers' privacy preferences can explain why they fail to adequately address their privacy concerns, leading to a higher level of consumption. Lastly, knowing the complexities of fixing this market, we investigate an alternative solution that could possibly replicate the planner solution without the need for intrusive government intervention.

Keywords – Privacy, Economics, Externalities, Pigouvian tax, Ignorance

Contents

List of Figures

List of Tables

1 Introduction

Online platforms have been subject to public scrutiny and criticism due to their allegedly dubious use of personal information and irresponsible distribution to third parties. Online platforms have the special feature of being multi-sided. That is, platforms that are characterised by two or more distinct sets of users who either consume or provide advertisements on one side of the market and sell or purchase personal data on the other (Agrawal and Fox, 2021). Hence, the vast majority of revenue from online platforms stems from advertising, where intimate data collection allows third parties to bid for presence in consumers' consciousness (Evans, 2009). This thesis focuses on the negative supply-side externality effect, where consumers disclose personal information about other individuals with similar characteristics when they consume services offered by an online platform (Acemoglu et al., 2019). It can be argued that if no externality exists in this market, then online platforms should be indifferent between a pay-for-privacy (PFP) option along with a free-version because they could extract all value from consumers in either case through price discrimination. However, online platforms usually provide their services for free which presents an argument that such platforms are advantaged by this socially harmful externality effect (Acemoglu et al., 2019). Besides this, it has been argued that technology-mediated business models may exploit the consumers' lack of knowledge as to how consumption of online services affects their privacy (Choi et al., 2019). The debate has shifted its locus of attention to safeguard the public's interest in preserving their own privacy. One example of a regulatory action to raise awareness about privacy risks can be seen in the European Union's General Data Protection Regulation (GDPR, 2021). In light of this, we discuss whether such regulation can educate consumers on the privacy risks of consuming online services, and investigate to what degree this may fix the inefficiencies in the market stemming from ignorance.

1.1 Research methodology

This master thesis makes use of existing literature on the topic of the economics of privacy to investigate the consequences of negative externalities and ignorance towards privacy preferences as two separate effects in the market for online platform services. Motivated by Acemoglu et al. (2019) and Choi et al. (2019), we first investigate the economic consequences of negative externalities, using economic theory to formulate the problem and interpret the results. Our model makes use of similar simplifying assumptions as Choi et al. (2019). Thus, we generalise the formulation by summing all online platforms into one monopolist platform that faces a homogeneous population. Furthermore, the social planner serves as a benchmark to evaluate the efficiency of the decentralised equilibrium. We use algebra to derive the key equations and express the relevant parameters and variables in functional forms to represent the market. Expressing the functions using parameters and variables might decrease some of the abstraction in our formulation, but in turn, provide simplicity for the reader to interpret the economic intuition of our investigation. Ignorance among consumers' privacy preferences and how to enhance consumers' control over personal data has been scrutinised over the years (see e.g., Colander (2017); van Ooijen and Vrabec (2019)). Ignorance, in this thesis, has been modelled by including a parameter to determine the perceived trade-off between online services and privacy. Although this parameter was included by brute force, the discussion provides justification for its inclusion.

With regards to the monopolist's profit function, we express the revenue from data collection in a reduced form which allows us to use economic intuition and prior research to identify an economic relationship between the variables. On a critical note, we want to make the reader aware that our formulation might be coloured by our and other secondary sources' prior understanding of the market, which might decrease the validity of our results and should be interpreted in light of this. Furthermore, we want to make it clear that the findings in this master thesis are not founded on empirical data and hence we are not implying any causal inferences. Rather, we provide an economic investigation to provide intuition and critical insights into the current practices and potential remedies in the market for online platform services.

To the best of our knowledge, we have identified a gap in the literature focusing on the consequences of both externalities and ignorance as two separate effects in this market. We acknowledge Acemoglu et al.'s (2019) analysis of users' undervaluation of own privacy, as well as Choi et al.'s (2019) attempt to investigate the fruitfulness of GDPR to educate consumers' awareness. However, a simple framework to understand the economic forces behind ignorance in online privacy decisions has to our knowledge not yet been investigated in this context. Combining the literature on the economics of privacy, welfare, and behavioural economics, we seek to establish an economic representation of the market for online platform services.

1.2 Research question

We want to provide an economic investigation on the consequences of negative externalities in the market for online platform services. Additionally, we investigate to what degree consumers' ignorance towards how consumption of service affects privacy can distort the market outcomes. This boils down our research question to:

"What are the consequences of externalities and ignorance in online platforms? An economic investigation of current practices and potential remedies".

1.3 Framework

First, we present an introduction to the field of information economics, which is the parent topic of the economics of privacy. Section 2 introduces our simple economic model which is used to investigate the consequences of negative externalities in a market for online platform services. This section formulates and derives the decentralised solution and the social planner's solution. Further, in section 3, we implement the planner solution and investigate whether Pigouvian taxation can replicate the socially optimal equilibrium. In section 4, we provide an extension of the model to investigate whether ignorance, as a separate effect, can be attributed to some of the market failures, where consumers are not acting in line with their privacy attitudes. Furthermore, we provide an economic justification for ignorance in this market, and present an additional extension of the model to see what level of ignorance is needed to replicate the social planner solution without the need for intrusive government regulation. Lastly, section 5 presents the conclusion which summarises the findings of this thesis.

1.4 Information economics

The field of information economics is not itself new, yet its implications have gained increasing relevance, especially for regulators, due to the emerging ecosystems using data as a vital input in value creation. “Knowledge is power” was the message from Stigler’s (1961, p. 213) Nobel prize-winning work on the economics of information and concluded that the "cold wind of ignorance is like subzero weather" (Stigler, 1961, p. 224). He argues with this statement that one can choose to do nothing and suffer the consequences of being ignorant to information, or one can adequately address the implications and live comfortably amid the phenomenon. The rise of data-intensive firms exemplifies Stigler’s relevance today, where data is a valued competitive advantage (Jones and Tonetti, 2020). As Acquisti et al. (2016) argue, information is not a single field, but an array of subfields, where privacy is only one among several interconnected fields that need to be accounted for to adequately address problems arising from information. Warren and Brandeis (1980) define privacy as the protection of someone’s personal space and their right to be left alone, whereas Westin (1967) defines it as the ability to control and safeguard personal information. Privacy, in this context, is defined as a "process of negotiation between public and private" (Acquisti et al. 2016, p. 40). When individuals choose to consume an online service, the online platforms will collect data as a by-product of consumption (Acemoglu et al., 2019), and sell this data to third parties. Hence, consumers’ privacy decrease in the amount of service consumed. Data is an unusual good due to its non-rivalrous and non-exhaustible nature, meaning that it can be consumed by several actors simultaneously without decreasing the remaining quantity (Jones and Tonetti, 2020). As a result, there are vast commercial gains that arise from data due to these properties, yet the consequences for consumers and regulators are not straightforward. This is partly because data is an extremely heterogeneous good, where its consequences depend on the context in which it is used, and partly because of its non-rival and non-exhaustible properties.

Advertisers may increase welfare by using data to offer personalised advertisement, as this might reduce searching- and matching costs (Aridor et al., 2020). However, online platforms can also gather information about their consumers' reservation prices (Acquisti et al., 2016). In doing so, they can offer personalised pricing strategies to extract all consumer surplus, which makes consumers worse off. Although price discrimination is an interesting facet of such markets, this thesis focuses rather on the two distinct effects found in externalities and ignorance, and how they are creating inefficiencies in this market.

Jones and Tonetti (2020) argue that online platforms have an obvious incentive to exploit personal information from consumers. In lack of government regulation, it can be argued that there are no real consequences of exploiting personal data, which incentivises online platforms to do whatever it takes to maximise profits from data collection. Research conducted by Acquisti and Grossklags (2005) and Sundar et al. (2013) shows that consumers of online services are aware of the privacy risks on the internet but disregard these threats in return for retail value and personalised services. Many scholars have argued that there is a dichotomy between privacy attitude and behaviour (see Acquisti (2004); Barth and de Jong (2017); Norberg et al. (2007)) which have been coined the "privacy paradox". We argue that understanding this dichotomy between attitudes and behaviour is important for regulators because this can be viewed as the first step towards fixing the problem of ignorance.

2 Model

In this section, we model the market for online platform services in presence of negative externalities using a simple economic model. First, we present a brief theoretical introduction about externalities to lay the foundations for the economic investigation of the market.

The following table (2.1) summarises the variables and parameters that will be used throughout this thesis.

Table 2.1: Symbols

Symbols	Description
s_i	Good: Service provided by the platform
p_i	Good: Privacy
C_i	Good: All other goods
Y_i	Income
S_{-i}	Total consumption of service, minus individual i
S	Total consumption of service
N	Number of consumers in the market
α	Parameter determining the marginal utility from service
β	Parameter determining the marginal utility from privacy
v	Price for service provided by the platform
γ	Marginal externality effect
δ	Parameter determining the marginal revenue from data collection
θ	Ignorance
τ	Tax
$H(S)$	Revenue from data collection
$c(S)$	Cost of producing service

2.1 Externalities

Externalities can be viewed as a by-product of either consuming or producing a good, imposing either a positive or negative value for others (Tietenberg and Lynne, 2018). Network effects are one example of a positive externality, where the value of a product or service increases with the number of people who participate. Another example that represents a negative externality is pollution. Alayo et al. (2017) consider externalities to be one of the greatest reasons for market failure, which has serious effects on society's overall welfare because consumers do not consider externalities when they maximise their utility. In a decentralised market with negative externalities, the price of a good is considered to be too low because the true cost of one's consumption is not internalised.

In our simple model, the platform collects, stores, and analyses user data to provide more personalised services, which in turn generates revenue through advertising or selling data to third parties. The prevalence of data mining and tracking technologies, such as cookies, intensifies the externality effects on the supply side, where the platform infers private information based on consumers with similar characteristics (Acemoglu et al., 2019). This externality effect explains how consumers impose additional costs on the economy without internalising the social damage they create.

2.2 The simple model

In going forward, we formulate and derive the maximisation problem for the consumers. The consumers' utility function is linear-quadratic to avoid corner solutions and ensures concavity:

$$u_i(s_i, p_i, C_i) = \alpha s_i - 0.5 s_i^2 + \beta p_i + C_i \tag{2.1}$$

Where $[\alpha, \beta] > 0$

There are N individuals who maximise their utility from consumption of service (s_i), privacy (p_i), and all other goods (C_i). α and β are parameters determining the marginal utility from consumption of s_i and p_i respectively. s_i can be characterised as a "no free disposal" good (Chellappa and Shivendu, 2010). The special feature of this is nonmonotonicity, where the utility first increases in s_i up until a certain level before it decreases in larger values. In our case, this can be viewed as the utility from consuming the online service provided by the platform, which is not necessarily increasing in quantity, but rather contains an intrinsic disutility at higher values due to individuals' privacy concerns (Chellappa and Shivendu, 2010).

The constrained maximisation problem is bounded by a monetary constraint

(2.2) and a privacy constraint (2.3):

$$Y_i = vs_i + C_i \tag{2.2}$$

$$p_i = K - s_i - \gamma S_{-i} \tag{2.3}$$

Where S_{-i} equals total consumption of service from all consumers except consumer i. Therefore $S_{-i} = S - s_i$ where S is total consumption of service from all individuals, thus $S = \sum_{i=1}^{N} s_i$.

Constraint (2.2) is the monetary constraint that prevents consumers from buying more than they can afford. Due to our static model, with only one period, the consumers find it optimal to spend all their money (Y_i). Additionally, v represents the unit price for the service provided by the platform.

Constraint (2.3) contains a constant K, which denotes the total endowment of privacy. Furthermore, p_i is a function of s_i and S_{-i}, meaning that privacy for individual i is negatively affected by own (s_i) and all other's (S_{-i}) consumption of services. This is a key element in our model, similar to Acemoglu et al. (2019) who argue that markets for online platform services may cause major inefficiencies or even break down in the presence of negative externalities. This, they argue, is because of the adverse effects that technology imposes on consumers with similar characteristics. They conclude that this mechanism will lower consumer i's utility, as the technology owned by the platform can infer private information through data mining, given that other people with similar characteristics share their data. In our model, this externality effect is captured by γ which imposes an extra negative effect on own privacy given other's consumption of service (S_{-i}).

2.3 The consumer problem

The consumers maximise the utility function subject to the two constraints:

$$\max_{s_i, p_i, C_i} \quad u_i$$
$$\text{subject to} \quad Y_i = vs_i + C_i \quad and \quad p_i = K - s_i - \gamma S_{-i}$$

Note that the consumers do not internalise the marginal externality effect (γ) imposed by others' consumption of service (S_{-i}) because the market level of consumption is outside the scope of consumers' control.

Further, we solve the consumers' maximisation problem by using the Lagrangian method:

$$\mathcal{L} = \alpha s_i - 0.5 s_i^2 + \beta p_i + C_i - \lambda_1(v s_i + C_i - Y_i) - \lambda_2(K - s_i - \gamma S_{-i} - p_i) \quad (2.4)$$

Yielding the following first-order conditions:

$$\frac{\partial \mathcal{L}}{\partial s_i} : \alpha - s_i - \lambda_1 v + \lambda_2 = 0 \quad (2.5)$$

$$\frac{\partial \mathcal{L}}{\partial p_i} : \beta + \lambda_2 = 0 \rightarrow \lambda_2 = -\beta \quad (2.6)$$

$$\frac{\partial \mathcal{L}}{\partial C_i} : 1 - \lambda_1 = 0 \rightarrow \lambda_1 = 1 \quad (2.7)$$

Substituting (2.6) and (2.7) into (2.5) gives:

$$\alpha - s_i - v - \beta = 0 \quad (2.8)$$

At this point, we assume that all consumers are homogeneous where α and β are the same for all individuals. Hence, $s_i = s$. Solving for s gives the equilibrium level of service in the decentralised market:

$$s = \alpha - v - \beta \quad (2.9)$$

Rewriting (2.9) to get the inverse demand function in terms of the price (v):

$$v = \alpha - s - \beta \quad (2.10)$$

The inverse demand function is determined by the marginal utility from consumption of service ($\alpha - s$) and the marginal utility from the loss of privacy (β) from consuming more s. The inverse demand function reflects the consumers' marginal willingness to pay for one extra unit of s. In principle, we can see that the consumers' marginal willingness to pay for service needs a high value of α to become positive. Thus, it is likely that v is negative. This can easily be seen because $\alpha < s + \beta$ is a possible scenario. The reason for this is that the utility function is nonmonotonic which reflects s as a no free disposal good, where a higher consumption of service implies a loss in privacy which enters negatively in the utility function.

2.4 The monopolist problem

The monopolist platform faces the inverse market demand denoted in equation (2.10) and sets a price v to maximise profits. The monopolist's profit function is given by:

$$\pi = vS - c(S) + H(S) \tag{2.11}$$

Again, consumers are homogeneous, hence $S = \sum s = Ns$

Note that the profit function shows a linear relationship between revenue and costs ($vS - c(S)$), where S is the total production of service offered by the firm. Additionally, the platform gains revenue through collecting and selling data to third parties, captured by the reduced form revenue function $H(S)$[1]. Hence, the monopolist faces a cost function (2.12) and a reduced form revenue function (2.13).

$$c(S) = cS + F \tag{2.12}$$

$$H(S) = \delta S - 0.5S^2 \tag{2.13}$$

The cost function consists of variable costs (cS) and fixed costs (F). Importantly, the variable costs (cS) are set to zero due to the negligible cost of producing and administering one extra unit of service from the platform (Chellappa and Shivendu, 2010). Therefore, we exclude cS from the cost function from now on.

[1] $h(x)$ describes the collection of data where x is data. Further, a function $g(S)$ captures data as a by-product of service. Lastly, $h(g(S))$ denotes the process of converting data into revenue, which we have dubbed $H(S)$.

The profit function contains the reduced form revenue function $H(S)$ which captures the additional revenue gains by collecting and selling data to third parties. This function describes how data is collected from consumers as a by-product of consumption (S) (Jones and Tonetti, 2020), and how the data is subsequently converted into revenue. Note that δ is the parameter determining the marginal revenue from selling the collected data. The function is linear-quadratic and concave which captures the diminishing returns to scale of revenues from data collection (Chellappa and Shivendu, 2010).

We start by inserting the inverse demand function from equation (2.10) from the decentralised market into the platform's profit function:

$$\pi = (\alpha - \beta - s)Ns - F + \delta Ns - 0.5(Ns)^2 \tag{2.14}$$

Further, we find the first-order condition:

$$\frac{\partial \pi}{\partial s} : \alpha N - \beta N - 2Ns + \delta N - N^2 s = 0 \tag{2.15}$$

Rearranging this and dividing by N gives the equilibrium level of consumption of service in the decentralised market, denoted s^D.

$$s^D = \frac{\alpha - \beta + \delta}{2 + N} \tag{2.16}$$

Equation (2.16) denotes the decentralised equilibrium of consumption of service. It is given by the parameter determining marginal utility from service (α), minus the marginal utility from privacy (β), plus the parameter determining the marginal revenue from data collection (δ), divided by $2 + N$. Note that γ does not enter in equation (2.16). Hence, the externality effect does not affect the level of consumption in the decentralised market.

Further, we derive the optimal price chosen by the monopolist by inserting the equilibrium level of service from equation (2.16) into the inverse demand function (2.10):

$$v^D = \alpha - \beta - \frac{\alpha - \beta + \delta}{2 + N} = \frac{\alpha(1 + N) - \beta(1 + N) - \delta}{2 + N} \tag{2.17}$$

2.4.1 Analysis of the price in the decentralised market

From equation (2.17), we can infer that larger values of δ and β negatively affect the monopolist price. This implies that if the platform values user data (δ) very much, and the consumers have a high marginal utility from privacy (β), then the platform could choose to subsidise the consumption of service because they receive more value from data collection by increasing consumption than they would through a positive price. This is the case in our model where a small value of α is not sufficient to compensate for the loss of privacy. More explicitly, we can see that when:

- $\alpha - \beta = \frac{\delta}{1+N}$, then $v^D = 0$,
- $\alpha - \beta > \frac{\delta}{1+N}$, then $v^D > 0$,
- $\alpha - \beta < \frac{\delta}{1+N}$, then $v^D < 0$.

The effect of $1 + N$ can be interpreted as the diminishing returns to scale of δ which enters in the $H(S)$ function from equation (2.13). The economic interpretation is that as N increases, the marginal valuation of data collection becomes smaller. That is, the value of α and β becomes more important parameters in determining the price as compared to δ when N becomes large.

In the case when $v^D = 0$, the platform provides service without a monetary transfer. In other words, the platform finds it optimal to provide the service for free because they will gain more value from data collection. Furthermore, it is realistic that online platforms implement a negative price when the marginal valuation of data collection is large. This is especially apparent for example in online banking platforms, where it is commonplace to offer monetary compensation for consumers to join their platform.

From this section, we derived the equilibrium level of service and price in the decentralised market. We further investigate whether this quantity and price are efficient by solving the social planner problem.

2.5 The social planner problem

Welfare economics seeks to solve the problems of what is the best level of consumption for society as a whole, all things considered (Sugden, 2013). The social planner problem can therefore be viewed as “the good God”, who cares about maximising both the consumers’ and the firms’ surplus, which constitutes the social welfare. If the social planner problem differs from the decentralised market, then there are reasons to consider government intervention.

2.5.1 Welfare implications of data collection

The formulation of the welfare maximisation problem is not straightforward. Following the definition of Sugden (2013), we must evaluate the contribution of data collection to society as a whole. Data collection may, for instance, decrease searching- and matching costs between firms and consumers by increasing the personalisation of advertisements, and therefore improve welfare (Hagiu and Jullien, 2011). On the other hand, data collection used for personalised advertising might be considered intrusive (Chellappa and Shivendu, 2010), increasing the actual or perceived risk of identity theft (Acquisti et al., 2016), or extract consumer surplus through price discrimination (Taylor, 2004). Besides the direct effects of engaging in data sharing, several scholars have argued that there are significant opportunity costs associated with keeping personal information private (Jones and Tonetti, 2020). For example, for health purposes, personal data could be used to diagnose diseases or predict flu outbreaks (Goel et al., 2010), or generally in product development by making existing products less labour intensive, for example like using artificial intelligence in transport (Jones and Tonetti, 2020).

It is clear that privacy is highly context-dependent and its effect on welfare is a complex task to gauge (Acquisti et al., 2016). However, in answering the question of whether to include $H(S)$ or not in the social planner problem, we must ask ourselves the question of whether this contributes to value creation for the platform and consequently for society. Choi et al. (2019) found that banning data collection in markets with strong negative externality effects represents a potential remedy to market failure. Instrumental to this argument is to gauge the magnitude of the negative externality effect. As have been supported in Choi et al. (2019) and Acemoglu et al. (2019), the social planner suggests a higher optimal quantum than the decentralised equilibrium if data collection is excluded, and therefore it might not be a better solution to ban data collection either, because this may also inefficient.

Hence, we cannot decide whether data collection is strictly good or bad for society. But if we, for the purpose of this discussion, assume that the platform belongs to a small open economy, which sells personal data to foreign actors, then the platform gains revenue from selling data, which of course creates value for the platform, and consequently for society. If we further assume that the marketing industry is a zero-sum game, then an increase in $H(S)$ can be considered "business stealing" (Mankiw and Whinston, 1986), and the platform grabs a larger part of this industry. It is worth noting that the platform can sell personalised data to third parties which may contribute to price discrimination. In that case, $H(S)$ would contribute to expanding the value of the respective industry, but the producer would extract all surplus (Grønn, 2008). This extraction of surplus, however, is not of interest to the social planner because we can regard this as a mere transfer of surplus from the consumers to the producer. In our formulation, $H(S)$ is a vital source of value creation for the online platform. Therefore, excluding this in the social planner problem would imply banning data collection as means of generating revenue, which we have argued might not be efficient either. We, therefore, choose to include $H(S)$ in this formulation, but we want to make the reader aware that it is not uniformly clear whether $H(S)$ is a positive or negative contributor to society as its consequences is highly dependent on the context it is used in.

2.5.2 Solving the social planner problem

From the discussion above, we choose to include $H(S)$ in the planner problem to maximise overall welfare. Hence, the value creation from the firm will be $H(S) - F$. To compare this value creation from the platform with the utility from consumers, we divide the value creation by N to distribute it equally back to society. To be precise, the explicit assumption is that the social planner redistributes the value creation from the platform's profit to consumers through lump-sum transfers. Hence, we include the additional monetary constraint:

$$C_i = \frac{\delta S - 0.5S^2 - F}{N} + Y_i \tag{2.18}$$

The social planner maximises the total welfare (W), which is the sum of utility from all individuals subject to the privacy constraint (2.2) and the additional monetary constraint (2.18). We formulate the social planner maximisation problem as follows:

$$\begin{aligned} &\max_{s_i,p_i,C_i} \quad W = \sum_{i=1}^{N} u_i \\ &\text{subject to} \quad p_i = K - s_i - \gamma S_{-i} \quad \textit{and} \quad C_i = \frac{\delta S - 0.5S^2 - F}{N} + Y_i \end{aligned}$$

Further, we solve the social planner problem using the Lagrangian method:

$$\begin{aligned} \mathcal{L} = &\sum_{i=1}^{N}(\alpha s_i - 0.5s_i^2 + \beta p_i + C_i) - \lambda_1 \sum_{i=1}^{N}(K - s_i - \gamma S_{-i} - p_i) \\ &- \lambda_2 \sum_{i=1}^{N}(\frac{\delta S - 0.5S^2 - F}{N} + Y_i - C_i) \end{aligned} \tag{2.19}$$

Where $\gamma S_{-i} = \gamma S - \gamma s_i$, which denotes the total externality effect imposed on every individual in the market.

Again, all consumers are homogeneous and hence, the planner will choose $s_i = s$, $p_i = p$, $C_i = C$, and $Y_i = Y$. We can therefore substitute the summation symbols with N and the total consumption from all consumers (S) equals Ns.

$$\begin{aligned}\mathcal{L} =& N(\alpha s - 0.5s^2 + \beta p + C) - \lambda_1 N(K - s - \gamma N s + \gamma s - p) \\ & - \lambda_2 N(\frac{\delta N s - 0.5(Ns)^2 - F}{N} + Y - C)\end{aligned} \tag{2.20}$$

First-order conditions:

$$\frac{\partial \mathcal{L}}{\partial s} : N\alpha - Ns + \lambda_1 N + \lambda_1 N^2 \gamma - \lambda_1 N \gamma - \lambda_2 N \delta + \lambda_2 N^2 s = 0 \tag{2.21}$$

$$\frac{\partial \mathcal{L}}{\partial p} : N\beta + N\lambda_1 = 0 \rightarrow \lambda_1 = -\beta \tag{2.22}$$

$$\frac{\partial \mathcal{L}}{\partial C} : N + \lambda_2 = 0 \rightarrow \lambda_2 = -1 \tag{2.23}$$

Inserting (2.22) and (2.23) into (2.21) and divide by N:

$$\alpha - s - \beta - \beta N \gamma + \beta \gamma + \delta - Ns = 0 \tag{2.24}$$

Rearranging this to get the socially optimal level of consumption of service, denoted s^S:

$$s^S = \frac{\alpha - \beta + \delta - \beta\gamma(N-1)}{1+N} \tag{2.25}$$

Now, we compare the level of consumption of service from the decentralised market with the socially optimal level:

$$(2.16) > (2.25)$$

$$s^D = \frac{\alpha - \beta + \delta}{2+N} > s^S = \frac{\alpha - \beta + \delta - \beta\gamma(N-1)}{1+N} \tag{2.26}$$

From the comparison in (2.26), we can infer that the socially optimal level of consumption of service yields a lower level than in the decentralised market. Note that the term $\alpha - \beta + \delta$ in the nominator is present in both cases, thus the only difference in the nominator between the two cases is the term $\beta\gamma(N-1)$. This term denotes the total externality effect on each consumer and explains how the equilibrium level of consumption in the decentralised market should, to obtain a socially optimal level, decrease in the total effect of the externality. This result aligns with our expectations because the consumers in the decentralised market maximise their utility without considering other's consumption of service, whereas the social planner considers the society as a whole, and thereby internalises the externality effect. N denotes the number of consumers in the market and is therefore assumed to be a large number. This implies that when $N \to \infty$, it is clear that the planner suggests a lower level of service than in the decentralised market.

However, if we consider the special case where $N = 1$, no externality effect exists in the market, but still the planner solution differs from the decentralised solution. As we can see from the comparison in (2.26), the planner solution is divided by $1 + N$ whereas the decentralised solution is divided by $2 + N$, implying that the planner will choose a higher level of consumption of service than in the decentralised market. The reason for this is that the monopolist will create a deadweight loss for society. This means that, in absence of a negative externality effect, the decentralised market actually provide less service than what the planner suggests.

So far, we can see that the price from the decentralised market does not lead to the socially optimal level of consumption of service (s^S). Hence, according to the planner, the monopolist platform should increase its price, given that $N > 1$, so that consumers end up at the socially optimal level of consumption suggested by the planner. This price is derived by inserting s^S from equation (2.25) into the inverse demand function from equation (2.10):

$$v^S = \alpha - \beta - \frac{\alpha - \beta + \delta - \beta\gamma(N-1)}{1+N} = \frac{N\alpha - N\beta - \delta + \beta\gamma(N-1)}{1+N} \tag{2.27}$$

Here, v^S is the price that obtains the socially optimal level of consumption of service. We can see that α and β affect the price in opposite directions, which reflects the planner's objective to maximise total welfare. The planner internalises the externality effect which is shown by the inclusion of the term

$\beta\gamma(N-1)$. Note here that the term $N-1$ is included because the negative externality effect on each consumer is imposed by all the other consumers in the market. Thus, the externality effect pushes the price upwards when N becomes large, which forces the price to reflect the true cost of consumption of service. Lastly, we can infer that the effect on price from δ is negligible when $N \to \infty$ because $N\alpha$ and $N\beta$ dominate the impact on the price. If, however, α and β are equally large and offset each other, then δ becomes more important in determining the price.

Naturally from the discussion above, we can see that the price from the decentralised market (v^D) and the price needed to reach the socially optimal level (v^S) also differ:

$$v^D = \frac{\alpha(1+N) - \beta(1+N) - \delta}{2+N} < v^S = \frac{N\alpha - N\beta - \delta + \beta\gamma(N-1)}{1+N} \quad (2.28)$$

The difference between the two prices is primarily found in the total externality effect $\beta\gamma(N-1)$. Again, if $N = 1$, we can see from the comparison of consumption in equation (2.26) that the socially optimal level is higher than the decentralised solution. Hence, we can infer that the price needed to reach the socially optimal level (v^S) would be lower if there is only one individual in the market.

So far, we have established that the decentralised quantity and price differs from the social planner solution. The special case when $N = 1$ provides and an argument that the monopolist platform might be advantaged by this socially harmful externality, because the decentralised solution suggests a lower optimal consumption of service compared to the planner solution when no externality effect exists.

3 Implementing the planner solution

In this section, we implement the findings from the planner solution in the previous section. Moreover, a discussion is included to critically evaluate the consequences of Pigouvian taxation in this market.

3.1 Pigouvian taxation

When there exist negative externalities in a market, the participants do not consider the full cost of consumption. A solution to this problem can be to "force" participants to consider the full costs of their actions (Pepall et al., 2014). This can be done by introducing a tax on the activities creating the externality. By doing so, the consumers will no longer underestimate the costs of their actions. Such taxes are called Pigouvian taxes and are levied by the government on actions that create socially harmful externalities. The concept of Pigouvian taxation is to declare the difference between social and private costs (Pigou, 1920). Hence, we want to derive a tax such that the monopolist platform chooses the optimal level of s equal to the socially optimal level (s^S). More formally, we introduce a tax τ in the profit function for the platform.

$$\pi = vS - c(S) + H(S), \tag{3.1}$$

where $c(S) = F + \tau S$

This ensures that the tax burden is levied on both the producer as a variable cost, as well as the consumers who need to pay for the burden they impose on society.

$$\pi = vS - F - \tau S + \delta S - 0.5S^2 \tag{3.2}$$

Further, we insert the inverse demand function from equation (2.10) into (3.2). Still, the consumers are homogeneous, and we, therefore, substitute S with Ns:

$$\pi = (\alpha - \beta - s)Ns - F - \tau Ns + \delta Ns - 0.5(Ns)^2 \tag{3.3}$$

Yielding the following first-order condition:

$$\frac{\partial \pi}{\partial s} : \alpha N - \beta N - 2Ns - \tau N + \delta N - N^2 s = 0 \tag{3.4}$$

Dividing by N and solve for τ gives:

$$\tau = \alpha - \beta - s(2 + N) + \delta \tag{3.5}$$

Further, we insert the socially optimal level of consumption of service from equation (2.25) for s in equation (3.5). Rearranging gives the tax which forces the monopolist to set the price that replicates the planner solution:

$$\tau = \alpha - \beta + \delta - \frac{\alpha - \beta + \delta - \beta\gamma(N-1)}{1+N}(2+N) \tag{3.6}$$

Simplifying this equation gives:

$$\tau = \frac{\beta - \alpha - \delta + \beta\gamma(N-1)(2+N)}{1+N} \tag{3.7}$$

Note that in equation (3.7), the tax is not equal to the negative externality effect because the platform's value creation is redistributed back to society. This means that the tax becomes smaller in δ, reflecting the additional value creation that each consumers' data is contributing to society. Another facet of this result is that the marginal valuation from data collection is diminishing in the total amount of users (N). This reflects the diminishing returns to scale from data collection, where the first unit of data is more valuable than the last one. If $\beta = 0$, ie., the consumers do not care about privacy at all, then the planner would promote consumption by subsidising consumers through a negative tax as this would be considered strictly welfare enhancing. However, this is not the case ($\beta > 0$), and we can see that β increases the tax, reflecting the planner's desire to maximise the overall welfare. Consumers still enjoy service, so the tax is decreasing in α as well as in the marginal revenue from data collection (δ).

In figure 3.1 below, we graph the effect of internalising the negative externality for the monopolist platform. Here, the marginal social benefit (MSB) is the demand curve, and the marginal private cost (MPC) is the supply curve, represented by the net marginal cost of producing the service.

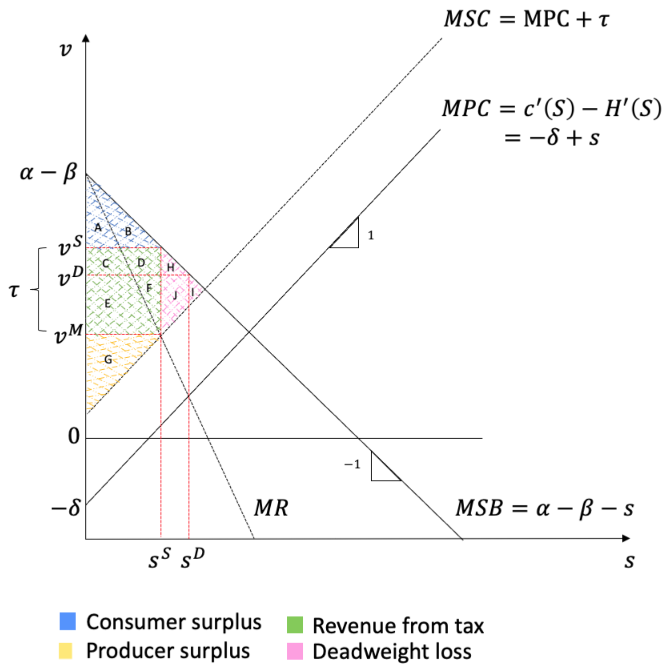

Figure 3.1: Pigouvian taxation

In the decentralised market, where the consumers do not consider the externality, the decentralised equilibrium is where the marginal revenue (MR) is equal to the MPC, leading the consumers to choose the level of consumption denoted s^D. Further, as corroborated in the model, adding a tax equal to τ will parallel shift the MPC with the amount of that tax, consequently yielding the marginal social cost (MSC) curve. Hence, this will force the consumers to consider the case where MR equals MSC, skewing the level of consumption to the socially optimal level, denoted s^S. This leads the consumers to pay a price equal to v^S, which is higher than the price in the decentralised market (v^D). Importantly, the monopolist platform will receive a price equal to v^M, which is lower than the original price (v^D).

From figure 3.1, we can see that the initial consumer surplus was given by the triangle $A + B + C + D + H$ and the producer surplus given by the triangle $E + F + G + I + J$. However, after introducing the tax τ, we can infer that the government earns tax revenue of the green area. Hence, both the consumer- and the producer surplus decreases, and therefore, they share the tax burden. Moreover, the size of the deadweight loss, as well as how the burden of a tax is distributed between producers and consumers, depends on the steepness of the supply and demand curve (Grønn, 2008). From our illustration, we can see that the producer bears more of the tax burden than the consumers, which is expected since the producer owns the technology that imposes the externality in the market initially.

3.1.1 Criticisms of Pigouvian taxation

On a critical note, Alcalde et al. (1999) argue that it is close to impossible for the regulators to have precise knowledge of utility and production functions, and therefore imposing a Pigouvian tax may have unpredictable consequences on consumption. Carlton and Loury (1980) further argue that levying taxes would not create an efficient outcome in the long run. This, they argue, is because the taxes are controlled specifically for individual firms, not for the industry as a whole. For the government to levy Pigouvian taxes, complete information is needed. This is almost never the case and might lead to inefficiencies in the market which eventually can lead to market failure. Similarly, William Baumol (1972) notes that measuring the social cost is close to impossible. He states that the social cost of any externality includes psychological and individual costs, which makes it extraordinarily difficult to measure the real costs. Therefore, according to Baumol (1972), the government can not find the optimal Pigouvian tax since it is not feasible to find the optimal level of consumption.

Furthermore, Ronald Coase (1960) criticised the need for intrusive government intervention such as Pigouvian taxation to fix the problem of externalities. Coase argued that if property rights are well defined and transaction costs sufficiently low, the problem of externalities will be internalised through negotiation without the need for government intervention. In our case, the property rights are well defined as the platform owns the technology that provides the service, and the consumers own money as well as privacy. Yet the market is still argued to be inefficient. Farrell (1987) argued that negotiation is not likely to result in efficient outcomes unless the parties have full information about each other. An observation made by Acemoglu et al. (2019), points out that the users of online platform services tend to share positive news more than negative ones. Therefore, according to Farrell (1987), negotiation will not be an efficient solution to the problem because the information is imperfect. Additionally, transaction costs of negotiation will most likely be high due to numerous consumers wanting to bargain to reach an efficient solution, hence we argue that the Coase theorem will fail.

3.1.2 Implications of taxing data collection

One issue with the current solution is that the tax burden is levied only on the production of service and not on the mere data collection that creates the negative externality effect. The problem is that online platforms base much of their value creation on the magnitude of their user base, which is challenging to measure the value of, and consequently also to tax. There is consensus among OECD and EU that taxation should be placed on profit where value is created (OECD, 2019), and therefore a clear definition of value in the context of digital business models is needed for it to be adopted in this case (Becker and Englisch, 2018). Failure to do so would imply major inefficiency costs due to administration and of course, lost revenue from taxation, in which case the economy could be better off without intervention.

Online platforms are distinct from conventional markets as they do not rely on physical presence, scale without mass, and intellectual property (Becker and Englisch, 2018). Consumers' interaction can be viewed as a part of online platforms' value creation and consequently profits and should therefore be taxed. On the other side, consumers are not themselves actively contributing to increasing the online platform's profits. They contribute to value creation merely by free will and "letting themselves be observed" (Becker and English 2018, p. 171). This added value can also be argued to be compensated through status and social interactions.

As established earlier in this thesis, data is a heterogeneous good which makes it very difficult to impose a standardised valuation framework for. It is challenging to know the value of data before it is realised in the market. For example, data about willingness to pay for a product is worth more to a platform than information about what an individual had for dinner yesterday. Data alone has no intrinsic value. The value from data is for example created by means of data mining, the process of finding patterns and inferences from data, which is done through intellectual property owned by the online platform. That is not to say that consumers can become suppliers of inputs to create value, but in such cases the consumers are normally compensated which appears as a cost for the firm, meaning that the firm is attributed for the additional value creation (Becker and Englisch, 2018). When online platforms offer consumers compensation in return for their data, for example through negative prices, then the value creation of selling data is attributed solely to the platform because consumers have already been compensated. Whether this compensation is sufficient will be discussed further in section 4.1, but as long as consumers choose by their own free will to consent to online terms and conditions of the platform, then we argue that the current formulation is fair.

3.2 Reluctance towards pay-for-privacy

From the discussion above, we can see the complexity of taxing the market for online platform services. However, this could be done substantially easier if the online platforms offered a PFP option in their menu. By doing so, the value of data collection would be realised in the market because the consumers could pay to opt out of it, consequently disclosing their reservation price for privacy. If all consumers choose this option, then all the revenue from the platform would come from the price, and hence the term $H(S)$ from our formulation would disappear. Therefore, the government could get substantially higher insight into this market and how online platforms value data, consequently allowing it to be regulated. As we have shown in equation (2.26), the monopolist platform chooses a higher level of consumption of service in the decentralised market than in the planner solution. Thus, the platform earns more profits if they can remain outside government intervention, and refraining to display the true price of their services represents a means to achieve this.

Hence, we argue that online platforms have an incentive to refrain from offering a PFP option because this would make it easier for the government to regulate the market, consequently lowering their profits. Tåg (2009) finds that a monopolist online platform will only employ a PFP option if the disutility that consumers get from advertisements is sufficiently high in relation to the advertisers' profit margins. This disutility can be translated in our model as the marginal valuation of privacy (β). Similar to the welfare implications of privacy, the consequences of offering a paid option to consumers may not necessarily be the best scenario either. For example, Tåg (2009) finds that the quantity of advertisement to consumers increases when platforms include a PFP option because the platform compares the revenue from paying consumers to the potential advertising revenue that can be earned by connecting those consumers to advertisers. If a large number of consumers would pay to preserve their data, the platform could arguably compensate for the loss of revenue gained through advertising by giving more advertisement exposure to those who remain in the data collection option. Introducing a PFP option might therefore risk only skewing the burden of advertisements over to those who remain in the data collection option, which might result in the welfare remaining unchanged.

So far, we have demonstrated how the consumers fail to consume the socially optimal level of service which can be attributed to the externality effect that the consumers fail to internalise. Moreover, we argue that there are vast complications of taxing data collection in the market for online platform services.

4 Extension of the model: Ignorance

In going forward, we investigate the degree to which ignorance among consumers can explain why they fail to adequately address their privacy concerns. Ignorance is a key issue in economics and can be defined as limitations in our knowledge (Colander, 2017). The consequence of ignorance is that consumers fail to act as economically rational agents when faced with privacy decisions. Acquisti (2004) argues that these privacy-related decisions are affected by asymmetric information and bounded rationality. The former explains how one agent can be under-informed in an economic transaction, and the latter how consumers fail to apply the existing information correctly.

First, we will introduce an extension of our model to investigate the economic implications of ignorance. Secondly, our results will be complemented by a critical discussion to investigate whether education can enhance consumers' awareness around privacy decisions, and consequently whether this is sufficient to fix the market inefficiencies stemming from ignorance.

4.1 Model with ignorance

The formulation is similar to the simple model but with the modification of $\hat{p_i}$ which represents the ignorance towards how consumers evaluate their privacy constraint. Thus, the utility function is now given by:

$$u_i(s_i, p_i, C_i) = \alpha s_i - 0.5 s_i^2 + \beta \hat{p_i} + C_i \tag{4.1}$$

Where $[\alpha, \beta] > 0$

This constrained maximisation problem is bounded by a monetary constraint and a privacy constraint:

$$Y_i = v s_i + C_i \tag{4.2}$$

$$\hat{p} = K - \theta s_i - \gamma S_{-i} \tag{4.3}$$

Here, θ denotes the level of understanding of how own privacy is negatively affected through higher levels of consumption of service (s_i). Still, the consumers are not able to internalise the externality effect. The current formulation allows us to evaluate the consequences of what would happen if consumers underestimate the consequences of how consumption of service compromises their privacy. If $\theta = 1$, then consumers are rational and therefore perfectly understand how consumption affects own privacy, which results in the equilibrium found in the previous section. Moreover, if $\theta < 1$, the consumers underestimate how their consumption affects their own privacy. Going forward, we investigate the consequences of such underestimation on price and quantity.

Again, we solve the consumer problem by using the Lagrangian method:

$$\mathcal{L} = \alpha s_i - 0.5 s_i^2 + \beta \hat{p}_i + C_i - \lambda_1(v s_i + C_i - Y_i) - \lambda_2(K - \theta s_i - \gamma S_{-i} - \hat{p}_i) \quad (4.4)$$

Yielding the following first-order conditions:

$$\frac{\partial \mathcal{L}}{\partial s_i} : \alpha - s_i - \lambda_1 v + \lambda_2 \theta = 0 \quad (4.5)$$

$$\frac{\partial \mathcal{L}}{\partial \hat{p}_i} : \beta + \lambda_2 = 0 \rightarrow \lambda_2 = -\beta \quad (4.6)$$

$$\frac{\partial \mathcal{L}}{\partial C_i} : 1 - \lambda_1 = 0 \rightarrow \lambda_1 = 1 \quad (4.7)$$

Substituting (4.6) and (4.7) into (4.5) gives:

$$\alpha - s_i - v - \beta\theta = 0 \quad (4.8)$$

Still, we assume that all consumers are homogeneous, and $s_i = s$. Solving for s_i gives the optimal level of consumption of the service for the consumers, denoted s^θ:

$$s^\theta = \alpha - v - \beta\theta \quad (4.9)$$

Now, we can see that when $\theta < 1$, the product of $\beta\theta$ will be lower than when $\theta = 1$. This implies that consumers will choose to consume more when they are ignorant compared to the case where they fully understand how their consumption of service affects their privacy.

Rewriting (4.9) to get the inverse demand function in terms of v, denoted v^θ:

$$v^\theta = \alpha - \beta\theta - s^\theta \tag{4.10}$$

The inverse demand function shows how the price (v^θ) increases as the consumers are ignorant, and consequently, how the ignorant consumers put a higher valuation on the service compared to the price in the decentralised market (v^D).

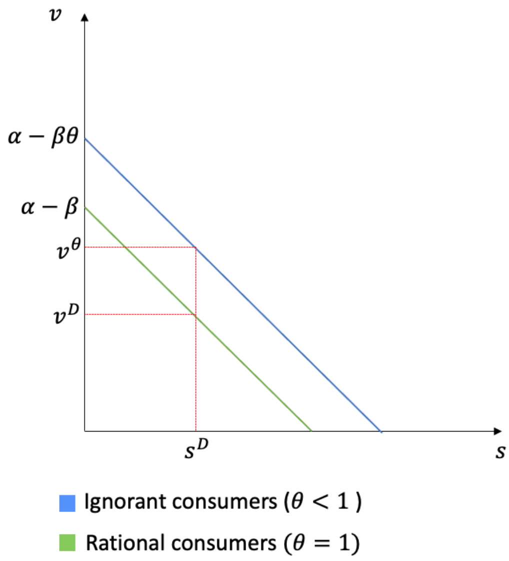

Figure 4.1: Price: Ignorant vs. rational consumers

In figure 4.1, the green line illustrates the inverse demand function from the decentralised market in our simple model, and the blue line illustrates the inverse demand function from the extension of the model with ignorance. This demonstrated how the price for ignorant consumers (v^θ) is higher compared to the decentralised market in our simple model (v^D). From the discussion in 3.1.3, we question whether consumers are adequately compensated for their data. From figure 4.1 one can argue that ignorant consumers pay a higher price for the same amount of service. Thus, we can make the argument that consumers' compensation decreases in lower levels of θ.

Again, we solve the profit maximisation problem for the platform using v^{θ}:

$$\pi = (\alpha - \beta\theta - s)sN - F + N\delta s - 0.5(Ns)^2 \tag{4.11}$$

Further, we solve for the equilibrium level of consumption of service chosen by the monopolist in the decentralised market as we did in section 2.4, now denoted s^{θ}:

$$s^{\theta} = \frac{\alpha - \beta\theta + \delta}{2 + N} \tag{4.12}$$

Comparing this result to the level of consumption from the decentralised market:

$$(4.12) > (2.16)$$

$$s^{\theta} = \frac{\alpha - \beta\theta + \delta}{2 + N} > s^{D} = \frac{\alpha - \beta + \delta}{2 + N}$$

Further, we derive the price:

$$v^{\theta} = \alpha - \beta\theta - \frac{\alpha - \beta + \delta}{2 + N} = \frac{\alpha(1 + N) - \beta\theta(1 + N) - \delta}{2 + N} \tag{4.13}$$

Here, we can also see that the price is higher than in the decentralised market:

$$(4.13) > (2.17)$$

$$v^{\theta} = \frac{\alpha(1 + N) - \beta\theta(1 + N) - \delta}{2 + N} > v^{D} = \frac{\alpha(1 + N) - \beta(1 + N) - \delta}{2 + N} \tag{4.14}$$

In this extension, it is clear that consumers choose to consume a higher quantity of the service at a higher price if they are ignorant ($\theta < 1$). This can be supported by Phelan et al. (2016), who argue that consumers dislike privacy intrusion, yet most individuals fail to protect their personal information, consequently disclosing too much information than what their attitude implies. When consumers make a privacy decision, they evaluate the risks against the benefits (Acquisti and Grossklags, 2005) to reach the decision that yields the highest net-benefit (White, 2004). This is exactly what we see in our extension, where the price for the service and the quantity increase as a direct consequence of insufficiently appraising how the consumption of service affects privacy.

4.2 Justification for ignorance

According to Acquisti et al. (2015), the ignorance among consumers makes it hard for individuals to understand their own privacy preferences. Further, Malhotra et al. (2004) describe how the privacy risk has a negative impact on the disclosure of personal information and, as was shown by Acquisti and Grossklags (2005), even the most privacy-conscious consumers fail to reserve their privacy in exchange for digital services. This demonstrates the degree of ignorance among consumers, which we argue is a result of not having the required information or not applying it adequately in order to make rational decisions.

4.2.1 Information asymmetries

Consumers' understanding of their privacy decisions can be severely hindered by asymmetric information (Acquisti et al., 2016), which is known to hamper economic efficiency and, as a consequence, welfare. Information asymmetries result in market failure where one party has more information than another in an economic transaction (Grønn, 2008). The information asymmetry, in this case, is that the producer has full knowledge about the effect that consumption has on consumers' privacy, whereas the consumers do not. In a study by Acquisti and Grossklags (2005) they show that consumers have a lack of knowledge about technological or legal forms of privacy protection which leads to misinterpretations of privacy violations and imprecise predictions of possible future risks. Moreover, the information asymmetries raise questions regarding consumers' optimal privacy decisions as rational individuals, and according to Acquisti et al. (2016) it is difficult for consumers to know when their data is collected, its purpose, and which consequences that come with it.

The consumers might not have the full knowledge of how much personal information is gathered by the online platform, leading to an imbalance of power in the transactions between the online platform and the consumers. By taking advantage of information asymmetries, the platform might be able to collect more personal information at a higher price than what consumers would accept if they were fully informed, which is exactly what we see in the extension of the model with ignorance. Moreover, and similar to the argument proposed by Jones and Tonetti (2020), if consumers believe that $H(S)$ is large, then this can be translated into $\theta > 1$. This results in consumers overvaluing the effect that consumption has on privacy and consequently consume less. On the other hand, if they believe that $H(S)$ is small, then $\theta < 1$, and consumers are undervaluing this effect and hence consume too much. However, it might also be the case that there is perfect information symmetry, but consumers still fail to apply all the information in their decisions, which is known as bounded rationality.

4.2.2 Bounded rationality

Bounded rationality is an inability to process and compare benefits against risks (Acquisti, 2004), and it limits an individual's ability to acquire, memorise and process all relevant information. In a privacy-sensitive situation, such as consenting to online terms and conditions, individuals might be limited by bounded rationality in the sense that they are unable to consider all relevant parameters. Kahneman and Tversky (1979) and later echoed by Acquisti (2004) claims that there are systematic psychological deviations from rationality that affect individuals in decision-making, which makes it difficult for consumers to protect their privacy. Therefore, people might believe that the benefit of using a service is higher than what they imagine the risk to be, and hence consume more of the service. The fact that consumers are ignorant, limited by asymmetric information, and bounded rationality, creates scenarios where attitudes and behaviour towards privacy decisions deviate over time.

4.2.3 The privacy paradox

Smith et al. (2011) argue that consumers are concerned about their privacy, including the distribution of their personal data to third parties. Barth and de Jong (2017) further argue that users of online services make relatively little effort to protect their privacy compared to the privacy concerns they are claiming to have. In a study by Norberg et al. (2007), they found that individuals disclose more information than they intended to. This leads to a dichotomy where there is a deviation between users' privacy attitudes and their actual behaviour, known as "the privacy paradox" (Norberg et al., 2007). This paradox can be partially explained by the mere failure of consumers' ability to value their own privacy (Lutz and Strathoff, 2013), or failure to understand their own utility functions. This paradox has raised the question of whether educating consumers on the true costs of consumption will make their actions start to align with their attitudes. However, for consumers to become less ignorant and for the market to reach the optimal solution, it can be questioned whether any government regulation should be introduced.

4.3 Remedies for ignorance

As mentioned above, ignorance among consumers may be a substantial contributor to market failure, and a central question is whether there is a need for a stricter government intervention to achieve higher awareness around privacy decisions.

Using government intervention to decrease market inefficiencies is a complex task due to the specific knowledge needed about what is driving consumers' privacy choices. This knowledge includes how behavioural and social forces make consumers disclose personal information initially (Acquisti et al., 2015). It can be argued that if the policymakers acquire enough knowledge about consumers' behaviour, the problem of ignorance may be reduced. In absence of sufficient understanding of behavioural and social forces, policymakers are likely to use instruments that are not accurate enough to empower consumers to act in line with their privacy attitudes.

An example of such a government intervention to increase awareness of privacy among consumers is the GDPR. When the GDPR was implemented in 2018, firms were required to do significant changes in how they collect, store, and manage personal data based on consent from the consumers (van Ooijen and Vrabec, 2019). One of the objectives was that firms should be clearer in their online terms and conditions, and to provide consumers with more control over their disclosed personal data. Thus, GDPR can be viewed as a proxy to evaluate whether government actions can raise privacy awareness in this market.

Recent studies have shown signs that the introduction of GDPR makes consumers more aware of their risks in online privacy decisions. For example, Aridor et al. (2020) used GDPR as a natural experiment to investigate through a diff-in-diff analysis the effects of GDPR as a whole. They found an approximate 12.5% decrease in total cookies accepted after the implementation of GDPR, pointing towards a strong effect for individuals with high privacy valuations. Moreover, Goldberg et al. (2019) found that the introduction of GDPR resulted in an 11.7% reduction of page visits as well as a 13.3% reduction in revenue from e-commerce. However, this is a fairly recent regulation and still lacks empirical evidence over time to draw reliable causal inferences from it. Additionally, the implementation of GDPR has been criticised for not adequately addressing the behavioural aspects of privacy (van Ooijen and Vrabec, 2019). Therefore, it can be questioned whether GDPR is decreasing the ignorance among consumers, which can be translated into the value of θ closer to 1 from our extension.

Furthermore, it can be argued that consumers consider the effort and loss of time in reading lengthy and complex terms and conditions to outweigh the perceived risk of disclosing personal information. This phenomenon is known as rational ignorance (Downs, 1957), where the individuals consequently evaluate the benefit of disclosing personal information to be higher than the risks, without actually knowing the consequences of disclosure. Similarly, van Ooijen and Vrabec (2019) argue that consumers rarely consider the consequences of consenting to online terms and conditions, which might provide another justification why individuals choose to consume more when they are ignorant in our model.

The government has the opportunity to force the online platforms to provide more accurate information about how data is collected and for what purpose it is used (Acquisti et al., 2016), which might result in more informed consumers, consequently increasing the value of θ closer to 1 in our model. On a critical note, however, it can be questioned whether the government actually has better knowledge of what is best for consumers, than the consumers themselves (Pelikán, 2010). After all, the true value of θ is not observable, and assuming that the government knows the true value of θ as opposed to consumers might lead to unintended consequences. In summary, we question whether government regulation is sufficient to enhance consumers' understanding of how their privacy is affected. This is because consumers will arguably choose to remain rationally ignorant since they believe that the perceived cost of privacy risk will not exceed the cost of reading online terms and conditions.

4.4 Paranoia as a remedy

So far, we have argued that people generally tend to undervalue the privacy costs of their online actions. Knowing the complexity of taxing data collection, nonetheless fixing the problem of ignorance, one might question whether taxing the market or educating consumers is optimal at all. If we consider the theory of the second-best (Lipsey and Lancaster, 1956), where we start in a market with presence of both externalities and ignorance, we can investigate at what level of θ that would make consumers find it optimal to replicate the social planner solution.

Again, remember equation (4.3) which shows that θ alters consumers' perception of how consumption of service affects their privacy. If $\theta < 1$ means that consumers are undervaluing how their own consumption affects privacy, and $\theta = 1$ implies rationality, then $\theta > 1$ can be considered paranoia. Paranoia can be provoked by the government through means of spreading information that makes consumers overestimate how consumption of online services negatively affects their privacy. This implies that the government, in theory, can increase consumers' perceived costs of consuming online services as an alternative to intrusive government intervention in this market. In the case of paranoia, consumers believe that the effect of consumption of service is in fact larger than its true value.

We investigate what the level of θ would be in order to reach the socially optimal level of consumption from equation (2.26), replicating the planner solution.

Using equation (4.12), and rearranging to get the expression of θ as a function of service gives:

$$\theta = \frac{\alpha + \delta - s^{\theta}(2+N)}{\beta} \tag{4.15}$$

Further, we insert the socially optimal level of consumption to find the value of θ which will yield this level. Hence, we insert s^S from equation (2.26) into (4.15) and solve for θ:

$$\theta = \frac{\alpha + \delta - \frac{\alpha - \beta + \delta - \beta\gamma(N-1)}{1+N}(2+N)}{\beta} \tag{4.16}$$

Simplifying this gives:

$$\theta = \frac{\beta(2+N) - \alpha - \delta + \beta\gamma(N-1)(2+N)}{\beta(1+N)} \tag{4.17}$$

As we can infer from equation (4.17), β, as well as the total externality effect, is multiplied by the population (N), whereas α and δ are not. Thus, it is likely that θ is larger than 1 in this expression. This can be seen by the fact that:

$$\alpha + \delta < \beta(2+N) + \beta\gamma(N-1)(2+N)$$

is a likely scenario, which means that the nominator is a positive number. Having established this, we can further see that:

$$\frac{\beta(2+N) + \beta\gamma(N-1)(2+N)}{\beta(1+N)} > 1$$

because

$$\beta(2+N) + \beta\gamma(N-1)(2+N) > \beta(1+N).$$

The level of θ from equation (4.17) will increase the perceived cost of privacy compared to its true value and, as a consequence, decrease the level of consumption of service down to where the externality effect is internalised, replicating the socially optimal level. Additionally, it can easily be seen that N affects θ through β. Thus, N is the most prominent parameter in determining the level of paranoia needed to reach the socially optimal level of s. When $N \to \infty$, the level of paranoia to reach the socially optimal level of consumption will also be higher. Conversely, if the market is small (N is a low number), then the externality effect is also smaller because there are less people in the market to impose externalities on. Thus, the level of θ needed to offset the negative externality effect also becomes smaller.

Moreover, having paranoia ($\theta > 1$) can be viewed in a similar vein as to how we taxed this market. Note that all parameters enter with the same signs as in the tax τ from equation 3.7. Because the marginal costs of producing an extra unit of service increase with the amount of the τ, the consumers no longer find it optimal to consume as much as before. The same mechanism can be seen here. Hence, paranoid consumers find it optimal to choose a lower level of consumption than in the decentralised solution because the loss of privacy from consumption of service is perceived to be too costly. However, note that this solution rests on the premise that $\beta > 0$. Hence, if $\beta = 0$, then $\theta = 0$, which means that the consumers do not care about privacy at all and therefore consume service to infinity. In this case, paranoia would have no effect on the quantity consumed. It, therefore, makes economical sense that β is included in the denominator to account for this mechanism. Thus, the government could, in theory, use paranoia as an instrument instead of tax to replicate the planner solution, given that consumers care about privacy.

In summary, we have established that in presence of externalities, it might actually be better for the economy that consumers are paranoid instead of fully rational. As mentioned above, the level of paranoia needed to replicate the planner solution is largely determined by the number of consumers on the platform (N). This reflects the argument that more consumers on the platform increases the negative externality effect due to the fact that the consumers disclose information about other consumers with similar characteristics (Acemoglu et al., 2019). Hence, a higher level of paranoia is needed to offset this effect if N is large. However, this argument completely disregards the importance of credibility for a government and might therefore be unrealistic in the real world. Still, this would solve the market inefficiencies in our model, and we, therefore, argue that it provides an interesting alternative solution to this market failure.

5 Conclusion

The heart of this thesis was to investigate the consequences of negative externalities and ignorance in consumers' privacy decisions, as two separate effects, in the market for online platform services. We have argued that consumers' privacy gets negatively affected by own, as well as other's consumption of online services. The latter represents the negative externality effect, which the consumers are aware of but do not control. From our simple model, we found that the social planner suggests a lower level of consumption than in the decentralised market. This is because the planner maximises the overall welfare for the society as a whole and therefore also considers the externality effect that the decentralised market fails to internalise. The fact that the decentralised solution differs from the planner solution is arguably by mere design, but it is supported by several scholars (see e.g., Acquisti et al. (2016); Choi et al. (2019)). Furthermore, we introduced Pigouvian taxation as a potential solution to "force" the market to internalise the negative externality effect. However, we argued in our critical discussion that the innate heterogeneity of data makes it virtually impossible for regulators to reach policies that restore first-best allocations because the regulators cannot get sufficient knowledge about supply and demand in this market. We also argue that the online platform will refrain from a PFP option because this would disclose how the platform values data collection, consequently making it easier to regulate.

As a critique of our model, using a homogeneous population somewhat contradicts our emphasis on looking at data as a highly heterogeneous and context-dependent good. If we introduced heterogeneity in our model, for example in privacy preferences, then different pricing strategies could be evaluated against the social planner solution and possibly yield a more realistic representation of the market. Hence, it might exist several decentralised equilibria that replicate the planner solution if heterogeneity in privacy preferences is included. Another critical aspect is the lack of network externalities in our model. It would be interesting to investigate whether this effect could offset the negative externalities, and perhaps reach a decentralised equilibrium closer to the socially optimal level.

In the extension of our model with ignorance, we investigated whether consumers ignorance towards how consumption of service affects their privacy can be attributed to some of the market failure. Here, we found that the consumers will end up paying a higher price for the same amount of service when they are ignorant (See figure 4.1). From this finding, we argue that ignorance is a source that makes consumers under-compensated in the transaction of service. Our results indicate that consumers choose to consume more of the service if they undervalue the effect that consumption of online services has on privacy. Moreover, we found that a government intervention to enhance consumers' awareness of privacy decisions is complex due to the apparent deviation between privacy attitudes and behaviour, and might not result in efficient market outcomes. Education might therefore play a significant role in the question of whether consumers are sufficiently compensated for their data. Importantly, our model shows that even if consumers are perfectly educated on their privacy risks, the market will still fail to reach the socially efficient equilibrium due to the presence of negative externalities. If we accept that the first-best equilibrium is unobtainable in this market, then the government could increase consumers' perceived cost of privacy to replicate socially efficient market outcomes. We found that the level of ignorance needed to replicate the planner solution is likely to be larger than one, which we have dubbed "paranoia". Hence, it might be the case that consumers end up internalising the externality effect themselves without the need for intrusive government regulation.

References

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdalgar, A. (2019). Too much data: Prices and inefficiencies in data markets. *National Bureau of Economic Research.*

Acquisti, A. (2004). Privacy in Electronic Commerce and the Economics of Immediate Gratification. *H. John Heinz III School of Public Policy and Management*, pages 21–29.

Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015). Privacy and human behavior in the age of information. *American Association for the Advancement of Science*, 347(6):509–514.

Acquisti, A. and Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy*, pages 24–30.

Acquisti, A., Taylor, C., and Wagman, L. (2016). The Economics of Privacy. *Journal of Economic Literature*, 54(2):442–492.

Agrawal, D. R. and Fox, W. F. (2021). Taxing goods and services in a digital era. *National tax journal*, 74(1):257–301.

Alayo, H., Rider, M. J., and Contreras, J. (2017). Economic externalities in transmission network expansion planning. *Energy Economics*, 68(C):109–115.

Alcalde, J., Corchon, L. C., and Moreno, B. (1999). Pigouvian taxes: A strategic approach. *Journal of Public Economic Theory*, 1(2):271–281.

Aridor, G., Che, Y.-K., and Salz, T. (2020). The Effect of Privacy Regulation on the Data Industry: Empirical Evidence from GDPR. *National Bureau of Economic Research, Inc.*

Barth, S. and de Jong, M. D. T. (2017). The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. *Telematics and Informatics*, 34(7):1038–1058.

Baumol, W. J. (1972). On Taxation and the Control of Externalities. *The American Economic Review*, 62(3):307–322.

Becker, J. and Englisch, J. (2018). Taxing Where Value is Created: What's 'User Involvement' Got to Do With It? *SSRN Electronic Journal*, 47(2).

Carlton, D. W. and Loury, G. C. (1980). The Limitations of Pigouvian Taxes as a Long-Run Remedy for Externalities. *The Quarterly Journal of Economics*, 95(3):559–566.

Chellappa, R. K. and Shivendu, S. (2010). Mechanism Design for "Free" but "No Free Disposal" Services: The Economics of Personalization Under Privacy Concerns. *Management Science*, 56(10):1766–1780.

Choi, J. P., Jeon, D.-S., and Kim, B.-C. (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics*, 173(6):113–124.

Coase, R. (1960). The Problem of Social Cost. *Journal of Law and Economics*, 3:1–44.

Colander, D. (2017). Ignorance and Economics. *Forum for Social Economics*, 46(2):139–144.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 65(2).

Evans, D. S. (2009). The Online Advertising Industry: Economics, Evolution, and Privacy. *The Journal of Economic Perspectives*, 23(3):37–60.

Farrell, J. (1987). Information and the Coase Theorem. *Journal of Economic Perspectives*, 1(2):113–129.

GDPR (2021). General Data Protection Regulation (GDPR).

Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., and Watts, D. J. (2010). Predicting Consumer Behavior with Web Search. *Proceedings of the National Academy of Sciences*, 107(41):17486–17490.

Goldberg, S. G., Johnson, G. A., and Shriver, S. K. (2019). Regulating Privacy Online: An Economic Evaluation of the GDPR. *NBER Working paper*.

Grønn, E. (2008). *Anvendt mikroøkonomi*. Cappelen Akademisk Forlag, 2 edition.

Hagiu, A. and Jullien, B. (2011). Why Do Intermediaries Divert Search? *RAND Journal of Economics*, 42(2):337–362.

Jones, C. I. and Tonetti, C. (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110(9):2819–2858.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *The Econometric Society*, 47(2):263–292.

Lipsey, R. G. and Lancaster, K. (1956). The General Theory of Second Best. *The Review of Economic Studies*, 24(1):11–32.

Lutz, C. and Strathoff, P. (2013). Privacy concerns and online behavior - not so paradoxical after all? {Viewing} the privacy paradox through different theoretical lenses. *SSRN Electronic Journal*, pages 81–99.

Malhotra, N., Sung, K. S., and Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, 15(4):336–355.

Mankiw, N. G. and Whinston, M. D. (1986). Free entry and social inefficiency. *The RAND Journal of Economics*, 17(1):48–58.

Norberg, P. A., Horne, D. R., and Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *The Journal of Consumer Affairs*, 41(1):100–126.

OECD (2019). Action 1 Tax Challenges Arising from Digitalisation.

Pelikán, P. (2010). The Government Economic Agenda in a Society of Unequally Rational Individuals. *International Review for Social Sciences*, 63(2):231–255.

Pepall, L., Richards, D., and Norman, G. (2014). *Industrial Organization Contemporary Theory and Empirical Applications*. John Wiley & Sons Inc, 5 edition.

Phelan, C., Lampe, C., and Resnick, P. (2016). It's creepy, but it doesn't bother me. *CHI Conference on Human Factors in Computing Systems*, pages 5240–5251.

Pigou, A. (1920). *The Economics of Welfare*. Palgrave Macmillan UK.

Smith, J. H., Dinev, T., and Xu, H. (2011). Information Privacy Research: An Interdisciplinary Review. *MIS Quarterly*, 35(4):989–1015.

Stigler, G. J. (1961). The Economics of Information. *Journal of Political Economy*, 69(3):213–225.

Sugden, R. (2013). The behavioural economist and the social planner: To whom should behavioural welfare economics be addressed? *Inquiry*, 56(5):519–538.

Sundar, S. S., Kang, H., Wu, M., Go, E., and Zhang, B. (2013). Unlocking the privacy paradox: Do cognitive heuristics hold the key. pages 811–816. Association for Computing Machinery.

Taylor, C. R. (2004). Consumer Privacy and the Market for Customer Information. *RAND Journal of Economics*, 35(4):631–650.

Tietenberg, T. H. and Lynne, L. (2018). *Environmental and Natural Resource*

Economics. Routledge, 11th edition.

Tåg, J. (2009). Paying to remove advertisements. *Information Economics and Policy*, 21(4):245–252.

van Ooijen, I. and Vrabec, H. U. (2019). Does the GDPR Enhance Consumers' Control over Personal Data? An Analysis from a Behavioural Perspective. *Journal of Consumer Policy volume*, 42(6221):91–107.

Warren, S. D. and Brandeis, L. D. (1980). The Right to Privacy. *Harvard Law Review*, 4(5):193–220.

Westin, A. F. (1967). *Privacy and freedom*. Atheneum.

White, T. B. (2004). Consumer Disclosure and Disclosure Avoidance: A Motivational Framework. *Journal of Consumer Psychology*, 14(1/2):41–51.