economies

Article

Data Protection, Cookie Consent, and Prices

Thomas Wein

Institute of Economics, Leuphana University of Lueneburg, 21335 Lueneburg, Germany; thomas.wein@leuphana.de

Abstract: A legislative process is currently ongoing in the European Union to supplement the 2018 General Data Protection Regulation regarding ePrivacy regulation. The supplement is intended to complete the European data protection policy in significant areas. One addition would be for service providers on the Internet, who currently obtain the consent of their users via an opt-out provision, to always provide a paid alternative without disclosing data. This procedure is essentially aimed at overcoming "cookie consent fatigue", which can be observed in many cases. A simple economic exchange model shows that users, as data subjects, are basically faced with the choice of paying a monetary price for a service that will also preserve their privacy or using Internet services "for free" while negating data privacy preferences. The individual demand for data privacy coincides with the socially optimal demand only if there is effective competition in the markets for data and Internet services and if users are sufficiently informed. In an online laboratory experiment with students of the Leuphana University of Lueneburg, a between-subjects design was applied in which the control group only had the option to either "pay" for the use of the artificial intelligence DeepL via cookies by surrendering data or to abstain from the service altogether, with the two treatment groups additionally given the option to use DeepL in exchange for a monetary fee so that privacy was not violated. To be tested was whether the "monetary price for privacy" option better reflected users' privacy preferences than the current cookie opt-out solution. The results show that it was much less common for DeepL to be remunerated with the disclosure of data and less common for DeepL to be waived entirely.

Keywords: data protection; cookie consent fatigue; privacy concerns; legal remedies

JEL Classification: K12; K24; L86

Citation: Wein, Thomas. 2022. Data Protection, Cookie Consent, and Prices. *Economies* 10: 307. https://doi.org/10.3390/economies10120307

Academic Editor: Ralf Fendel

Received: 22 September 2022
Accepted: 25 November 2022
Published: 1 December 2022

1. Introduction

Under the legislative process currently being pursued in the European Union for ePrivacy regulation, service providers on the Internet who wish to obtain their users' consent for the use of data by means of an opt-out rule must always have a paid alternative available. Users must be able to use the Internet provider's service without disclosing their personal data by choosing to pay a fee. This measure is intended to overcome "cookie consent fatigue" (Burgess 2018; Utz et al. 2019). Cookie consent fatigue occurs when users simply agree to the data privacy disclosure when asked for their consent, without reading the disclosure or thinking about what it means. We ask the question, "Does the planned regulation help ensure the desired individual level of privacy or data protection?"

From an economic point of view, data protection is focused on how personal data are collected, processed, and transferred. It is about protection of a person's private information and their right to be left alone (Acquisti et al. 2016). It is about protecting against intrusion into personality. In particular, it is about protecting individuals from the state interfering in their lives and restricting their ability to make decisions. When information about an individual becomes known, the state could use it to restrict the freedom of individuals (Solove 2006).

Acquisti (2010) and Larouche et al. (2016) see a variety of potential benefits for consumers when they disclose data: (a) recommendations for products and regional suppliers that have characteristics that fit well with consumer preferences; (b) consumers have to search less intensively for products and suppliers, which reduces their search costs; (c) personalized products are suggested, in extreme cases generated via the 3D printing process; (d) unwanted, non-matching advertising is avoided; (e) in social networks, communication with other participants in the network is made possible; (f) free content on the network is offered to the user; (g) ratings on the Internet can improve one's social status; and (h) users may be able to sell information about themselves. It can therefore be seen that the disclosure of data can very well be in the interest of an individual.

The disclosure of data can also lead to disadvantages for consumers: (a) information intermediaries may expose consumers to advertisements they do not desire to receive and that they would not otherwise receive without disclosing their data; (b) consumers are directed via targeted advertisements to products associated with excessive prices; (c) consumers may suffer harm when providers pressure them into consumer behavior that is disadvantageous; and (d) data may be used in an unpredictable way in the future, for example in the form of loss of reputation. If personal data is combined in an unpredictable way with other publicly available data, the negative effect may be compounded, for example: (a) personal data can be misused or stolen; (b) individuals may find themselves psychologically affected by the fact that information about them is in the hands of others; (c) information may be used to a person's detriment in the future when it is significantly more valuable; and (d) violations of privacy may result in material losses (e.g., higher prices), immaterial harms (e.g., anxiety), minor detriments (e.g., spam messages), or serious negative consequences (e.g., refusal of a mortgage in the case of identity theft or the denial of health insurance coverage). For more examples, see Acquisti (2010) and Larouche et al. (2016).

It is often feared that many individuals unintentionally disclose their data, although they claim in surveys that they do not want to disclose data—a "privacy paradox" (Barth and De Jong 2017). From the perspective of behavioral economics, the behaviors can be explained by (a) underestimating the long-term consequences of data pricing; (b) inappropriately preferring the current benefits of free use over the long-term disadvantages of loss of privacy; (c) that the way users respond to a data disclosure may mistakenly lead to wrong outcomes; or (d) inadvertently making wrong decisions. For example, selecting an opt-out box instead of opt-in box (Acquisti et al. 2015; Barth and De Jong 2017; Larouche et al. 2016; and basically rejecting Solove 2021).

The way the interests of data owners and data subjects interact in markets has been studied by economists for many years. Posner (1978, 1981) worried that privacy protection prevents the efficient exchange of information. For example, when unqualified applicants to a job are allowed to withhold characteristics relevant to the employer and are hired, the firm's productivity may be negatively impacted. At times, qualified applicants may not even be considered. Hiring unqualified applicants results in a firm's costs increase, and higher prices for consumers. Stigler (1980) takes this thesis even further by stating government interventions to protect poorly qualified applicants from revealing their data are at best ineffective and at worst harmful. They are ineffective because qualified applicants disclose information and unqualified ones "reveal" themselves as unqualified by withholding information. Harm occurs when the intervention is useless and generates costs. According to the work of Coase (1960), Laudon (1996), and Varian (2002), individual bargaining between both market parties discloses the efficient amount of information if the property rights to the data are clearly defined. Hermalin and Katz (2006) show that data disclosure can be associated with negative consequences. When it comes to insurance, data disclosure may result in being unable to insure oneself as a policyholder with a high loss potential, or being able to insure oneself only very expensively. To ensure ex ante insurability of all risks, policyholders may prefer a ban on disclosure.

In this paper, we restrict ourselves to looking at private consumers who might disclose data about themselves and producers who might use the collected data to produce Internet

services. We ask whether the planned mandatory provision of Internet services, even with a fee, will lead to better consideration of privacy preferences. A simple economic exchange model shows that users, as data subjects, are faced with the choice of paying a monetary price for a service and preserving their privacy or putting privacy preferences second and using Internet services "free of charge". The individual demand for data privacy coincides with the socially optimal demand only if there is competition in the markets for data and Internet services, and users are sufficiently informed. To investigate this matter further, we conducted an online laboratory experiment with students from Leuphana University concerning the use of the artificial intelligence DeepL. The Control Group can either forgo the service DeepL by opting out via cookies or consume the service by agreeing to lower privacy via cookies. The two Treatment Groups had the additional option of choosing more privacy and paying a direct fee for DeepL; either they revealed a self-imposed willingness to pay greater than zero or they voted in a student ballot for a campus version of DeepL with an annual fee of EUR 10. The additional options for Treatment Groups to use DeepL with good privacy for a monetary price should better reflect the privacy preferences of users by accepting cookies with low privacy less frequently or doing without DeepL altogether. By specifying a concrete (market) price for the campus license (Treatment Group 2), privacy preferences should be taken into account even better than when stating willingness to pay (Treatment Group 1) because of observing a realistic price signal. Participants who expressed a sufficient willingness to pay for DeepL could use the paid version of DeepL currently available to the public for one year at a price of EUR 10; the difference from the higher actual price would have been reimbursed from university budget funds. In this respect, there was an ex ante incentive to seriously pursue participation in the experiment. The use of this between-subjects design is permissible because the three randomly selected groups differ little in their characteristics, views, and preferences.

Section 2 provides a brief overview of the current legal framework and describes the plans for ePrivacy regulation. Furthermore, the section summarizes available empirical evidence on the previous ePrivacy Directives and the General Data Protection Regulation, including reporting existing field experiments. Section 3 presents a market model of the individually and socially optimal demand for privacy. The experiment with Lueneburg students and its results are outlined in the Section 4. A summary and conclusion can be found in the Section 5.

2. General Data Protection Regulation and ePrivacy Regulation

The relatively new European General Data Protection Regulation (GDPR) can be seen as the current gold standard data privacy legislation (Buttarelli 2016). The regulation became applicable law in May of 2018 (De Hert and Papakonstantinou 2016). Building on the understanding that privacy is a fundamental human right in the EU (Article 8 European Convention of Human Rights; Article 7 European Charter of Fundamental Rights), the GDPR strengthens traditional privacy principles relating to the processing of personal data. For instance, it covers data processing lawfulness, purpose, and storage limitation, as well as data minimization (Article 5 GDPR).

Correspondingly, data processing is only lawful if prerequisites are fulfilled, such as obtaining the data subject's consent, the existence of a contract or another legal obligation, or if greater public interests prevail (Article 6 GDPR). The GDPR further provides data subjects with extensive and partially novel rights, such as the right to rectification, the right to erasure, the right to data portability, and the right to object (Articles 16, 17, 20, and 21 GDPR). Ultimately, rigorous liability and compensation payments, as well as administrative fines for unlawful conduct by data holders, reinforces the privacy rights of data subjects (Articles 82 and 83 GDPR). The standard case of consent sets legally significant hurdles (Buchner 2020, paras. 9–81). Consent depends on the prior agreement of the data subject, as subsequent consent is not sufficient. Consent must be conscious, voluntary, and informed. Voluntariness is in question if de facto coercion is exercised, if there is a clear power imbalance, or if an unreasonable "take it or leave it" decision exists. The data

subject must be able to make an informed decision before giving consent (i.e., the required information must be sufficient, clear, and comprehensible to obtain informed consent). The consent must be sufficiently specific about the disclosure of the data concerned.

Along with the GDPR, additional data protection regulations apply, which are either anchored in the national Telemedia Act (TMG) or the Telecommunications Act (TKG), or European e-Privacy regulations, which are still based on directives from 2002 and 2009 and are set to be replaced shortly by the new ePrivacy regulation.

In the past, the European Union adopted two ePrivacy Directives. The first was in 2002, and the second in 2009. The latter included the obligation to obtain consent for cookies under an opt-in regime (González et al. 2020). As in the GDPR, the processing of personal data under sections 14 f of the TMG apply the principles of necessity and reservation of consent. According to Section 15 III of the TMG, cookies or comparable technologies that can record online usage profiles may be used if subscribers have consented on the basis of clear and comprehensive information, subject to the limits of civil law control of general terms and conditions, such as the inadmissibility of unreasonable disadvantage (Buchner 2020, paras. 161–174).

The new ePrivacy regulation, which will replace the previous two directives, is intended to provide greater protection for the rights of all users of electronic communications (SMS, e-mails, Facebook messages, social networks, etc.) (González et al. 2020; Council of the European Union 2021). Continuing in the spirit of the GDPR, gaining consent prior to the collection of data remains central. Specifically, the following shall apply:

- The regulation should apply to electronic communications of publicly available services and networks, including metadata (location, time, and data about recipients).
- The regulation protects European users, regardless of whether the service provider is located inside or outside the EU.
- The council proposes that explicit consent to cookies should only be valid if the user has the alternative option of paid use of the services and networks without cookie trackers.
- Users should be given the opportunity to set default settings for cookies—whitelists—via their browser settings, as well as to be able to easily change or revoke them. Whitelists of this kind are intended to counteract the problem that many users are overwhelmed by the large number of queries about cookie settings and therefore simply consent without recognizing the consequences of their consent.

According to the future ePrivacy regulation, consent should also be dispensable if there is no, or only very minor, threat of intrusion into privacy (Buchner 2020, para. 175).

The consequences of European data protection regulation were recently empirically investigated. From the analysis of almost 10,000 field studies on advertising campaigns from 2001 to 2008 in five European countries, Goldfarb and Tucker (2011) concluded 65 percent lower advertising effectiveness (less web bugs, cookies, and clickstream data) in online commerce. This trend is supposedly related to the e-Privacy Directives from the 2000s and its implementation in the member states.

There are a number of studies regarding the effects of the new GDPR, enacted EU-wide in May 2018. Goldberg et al. (2019) used data from the Adobe Analytics platform of 1508 companies to look at how the GDPR has affected salient metrics for European websites. Using a difference-in-differences approach, the new regulation appears to cause about a 10 percent decline in pages viewed, actual time of use, transactions made, and revenue generated. Johnson et al. (2020) extended the analysis by compiling a panel of over 27,000 websites, drawn from the 2000 most important Internet players from the EU, the USA, Canada, and worldwide. A special information technology tool was used to capture the relationship between internet players and web technology vendors, predominantly concerning advertising, web hosting, audience measurement, and social media. They looked at data from just before the GDPR came into force in May 2018, compared to the end of 2018 when the regulation was in full effect. In the short term, the market shares of the small web technology vendors fall in favor of the large ones such as Google or Facebook, which can be interpreted as causal due to the use of the difference-in-differences approach.

They attributed this short-term concentration movement to the fact that large vendors were more likely to be able to provide guarantees that they could meet the new data protection requirements. After a few months, the shift seemed to disappear. Building on the previous data set and using similar methodology, Goldberg et al. (2021) found that after the GDPR took effect, page views fell by an average of 11.7 percent. Recorded revenue also fell by 13.3 percent. Just under 10 percent of users refused consent, which explained the 9.4 percent effect on page views and 7.6 percent effect on revenue. The large remainder came at the expense of the GDPR in the form of more difficult marketing activities, whether among consenting or non-consenting users.

Lefrere et al. (2020) compared websites in the EU versus those from non-EU countries in terms of website content. According to their estimates, the introduction of GDPR in May 2018 did not seem to have a negative impact on the content of websites. Peukert et al. (2020) observed over 110,000 websites over 18 months before and after the GDPR introduction. On one hand, there were more informative privacy policies, but on the other hand, Google's market power increased. Jia et al. (2018) concluded from their data that in the short term following the tightening of data protection in Europe in 2018, investments in young, European tech companies declined compared to investments in the United States. Their follow-up paper (Jia et al. 2021) showed similar findings. Using a difference-in-differences approach, Aridor et al. (2020) used data from European and U.S. online travel agencies from the first eight months of 2018 to find the following effects from the newly introduced GDPR opt-out rule. Approximately one-eighth of users took the opt-out option, but the remaining users stayed longer on the website. Internet providers could therefore track users for a longer period of time, which they described as an eight percent higher trackability.

Utz et al. (2019) conducted field experiments on how the frequently observed cookie consent fatigue could be reduced by designing websites differently. For about 1000 consent variants on currently existing websites of a German e-commerce provider, more than 80,000 real website visitors were confronted with different forms of consent. Experiment 1 looked at whether the position of the consent notice on the screen influenced the decision (14,135 visitors). Experiment 2 asked whether the number of choices and the graphical highlighting of answers (nudging) had an impact (36,530 visitors). Experiment 3 tested whether the link to a privacy policy or the use of non-technical language, such as "this website collects your data", was relevant to 32,225 visitors. Approximately 100 participants took part in an online follow-up survey. According to the descriptive results, users were more likely to respond to a notice at the bottom left of the screen. They were more willing to accept cookies if they had the choice to accept or not accept them, instead of just agreeing to them, and they responded strongly to nudging.

Machuletz and Böhme (2019) investigated how 150 mostly first-year undergraduate computer science students at the universities of Muenster and Innsbruck responded to different consent dialogues and how they evaluated their decision afterwards. Without disclosing the goal of the study, participants were given the task of searching for a flight with departure and destination locations and dates. The search engine website had different, randomly assigned privacy notices that served as the lab experiment. Treatment Group 1 (the "deception variant") received a comprehensive text with three levels of data disclosure to choose from. If one chose the graphically highlighted button, all three levels were selected. Treatment Group 2 received only one level to choose from, but with the consent button highlighted. The Control Group could select the three options individually and the confirmation button was not highlighted. Users voted for more data collection purposes with highlighted buttons when there were multiple options to choose from, and much more so than in the work of Utz et al. (2019). The follow-up survey strongly suggested that participants were deceived in both treatments. The number of areas for which consent can be given had no significant effect.

3. Data Protection in Markets

Adapting from the textbook version of general equilibrium in Nechyba (2018), the following model shows how an individual, W, offers an Internet service based on data as well as providing the data itself, as well as how it acts in a profit-maximizing or utility-maximizing manner in competition. It is a matter of how much data the individual voluntarily discloses. In other words, how much privacy would the individual give up, and how much data would not be disclosed? The model shows it corresponds to the optimum level of data protection in society as a whole.

The individual W produces only one service, x She provides information in connection with the Internet and with the help of a search engine, or she sells a good or service by using the Internet. In either case, provision of information is always present. W thus becomes a data holder. Additionally, the individual W also provides data about herself, i.e., she gives away information about herself, the data subject. In extreme cases, she becomes a "glass human being" by providing too much information.

Figure 1 shows how the data owner W uses data d to provide service x. The production function $x = Ad^{\beta}$ with A as a constant and $0 < \beta < 1$ (production function is concave) applies to her. Consequently, W produces with positive but diminishing marginal returns if all other input factors such as labor and capital are constant.[1] In other words, more data from individuals help to provide x, but the increase in x decreases as data availability increases.

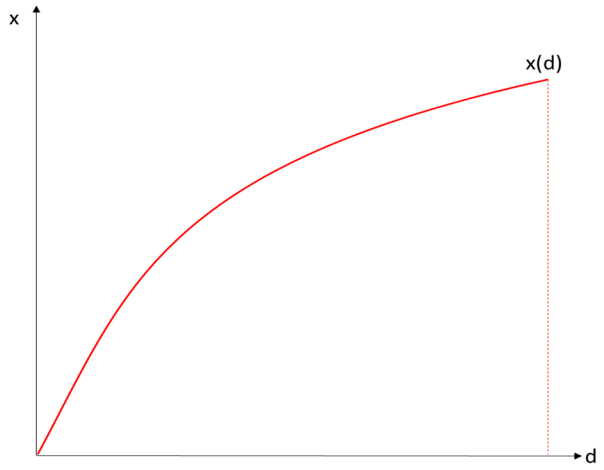

Figure 1. Production function of a data owner.

The data subject derives utility from the service x and disutility from the disclosure of its own data to third parties. In Figure 2, let the point S be given with the consumed amount of x_S and the disclosed amount of data d_S. Starting from d_S, a disclosure of further information in the amount of +d would mean a utility loss for the data subject, which would just be compensated by the increase in x by the amount of x`. S and S` lie on the same indifference curve I_1. Moving from S`` to S```, i.e., by definition the same amount of data disclosure +d, the individual "demands" a significantly larger amount of x to be compensated for the loss of privacy. In other words, the more individual data already disclosed, the more severe the resulting inconvenience. Only higher amounts of x can "compensate" for the loss of privacy. At the disclosure level D, at d = D, the data subject has disclosed all her data, and the level of privacy is zero. The data set D thus becomes the endowment point for the individual, which can be maximally disclosed. For example, the data subject's utility function can be written as $x^{\alpha}(D - d)^{(1-\alpha)}$ (Cobb–Douglas utility function with $0 < \alpha < 1$). In Figure 2, the point V compared to S' for a given price of data d means a higher availability of the service x, and consequently a higher indifference curve I_2 is obtained. Indifference curves that run to the upper left, such as I_3, are associated with even higher utility.

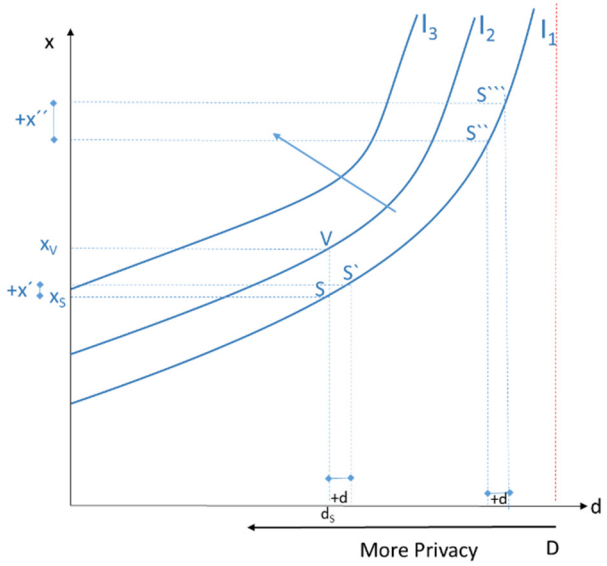

Figure 2. Indifference curves of a data subject.

Considering the data owner and data subject as one unit, the highest possible utility on the indifference curve I_3 would be reached at point C* for a given production function x(d) (Figure 3). Consequently, the data quantity d* would be revealed and not the maximum quantity D. Of the service x, x* is produced and consumed, but not up to x_D. In this respect, the non-consumed difference between x_D and x* can be interpreted as the "price" for preserving (partial) privacy, D-d*. The difference D-d* would thus be the optimal privacy for society as a whole if the same production and utility functions apply to all individuals; however, the data d* are disclosed in the optimum circumstance (see Appendix A).

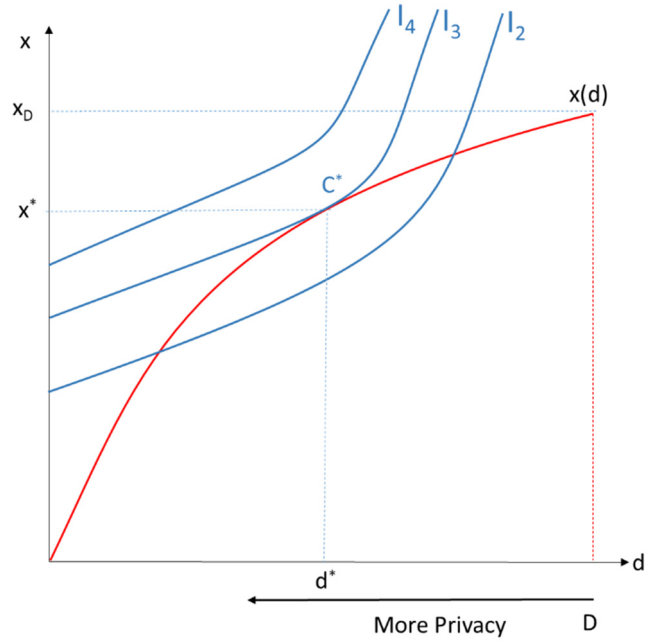

Figure 3. Welfare-optimal data protection.

If we return to separate individuals for data owners and data subjects, owners face the competitive goods price p for their provided goods. Data subjects can obtain the competitive factor price t for their data, and we can show that utility-maximizing or profit-maximizing individuals experience socially optimal privacy. Factor prices for data are either explicit prices (the data holder pays a price for data collected from the data subject) or

are implicit (the data subjects forgo monetary compensation for loss of privacy in exchange for quantities of x for free). Prices for the use of Internet services correspond either to direct monetary amounts (a "cash price") or to the utility losses of limited privacy if no monetary payment is made.

All combinations of quantities of data d used, for which the constant factor price t has to be paid, and different levels of services produced, which are sold at the constant price p and lead to the same profit level π, represent isoprofit lines (Figure 4). Consequently, in general,

$$\pi = px - td$$

$$x = \frac{\pi}{p} + \frac{t}{p}d.$$

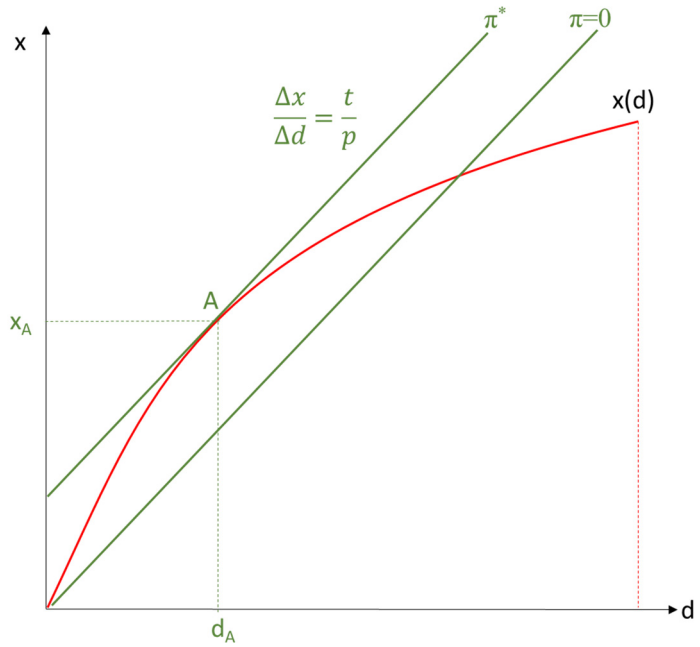

Figure 4. Data holder and optimal data demand.

The slope of the isoprofit line $\left(\frac{\Delta x}{\Delta d}\right)$ is thus equal to $\frac{t}{p}$. Constant prices imply competition in the service market for x and in the factor market for d. The isoprofit line through the origin corresponds to zero profit. Only goods along the production function x(d) can be produced at most. The data holder achieves highest possible profit in A, such that d_A is demanded as data disclosure and quantity x_A is supplied; the profit achieved is π^*.

Given commodity and factor prices p and t or isoprofit line π', the data subject chooses the highest possible indifference curve I_4 with point B, and thus the utility-maximizing data disclosure d_B or the optimal level of privacy D-d_B (Figure 5). The utility-maximizing amount of consumption occurs at x_B.

If there were no effective data protection, data owners could obtain data for free (t = 0), and isoprofit lines run horizontally to the abscissa (Figure 6). The highest possible isoprofit line π^{**} is reached at D where the data subject discloses too much of her data, more precisely, all her data (privacy D-d = 0), and consumes too many data-driven goods (x_D). The data subject then achieves a lower utility level ($I_1 < I_4$) than she would with an effective data protection policy in place.

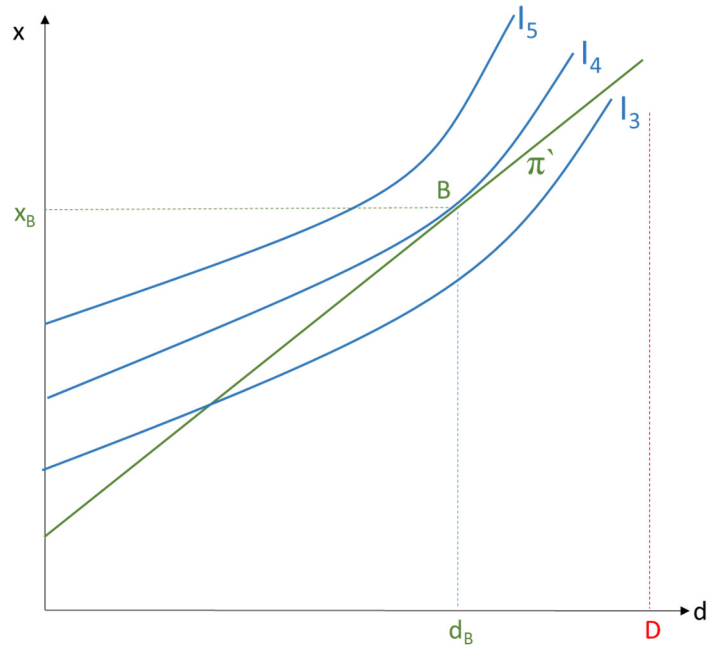

Figure 5. Data subject and optimal data supply.

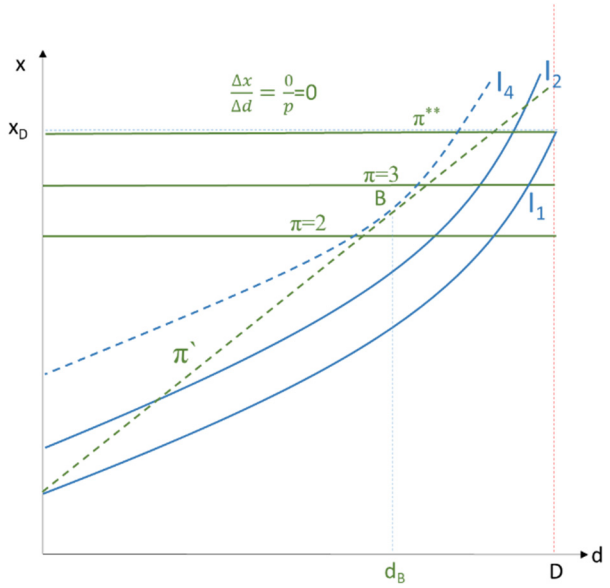

Figure 6. No effective data protection.

For equilibrium on the data market and thus for the realization of the optimal data protection level (D-d*) or the optimal amount of use of the data in B (d = d*), the data market must also be cleared. In Figure 7, more data would be offered than demanded, $d_B > d_A$. The data usage price t would have to decrease toward the equilibrium. At the same time, the usage price p would be too low because there would be excess demand for x ($x_B > x_A$). According to Equation (2), the slope of the isoprofit line t/p would have to fall until demand equals supply in both markets (t*/p*), which would be the case at point C** in Figure 8. As shown in Appendix A, Point C* in Figure 3 and Point C** in Figure 8 coincide. In the data market with utility-maximizing data subjects, the optimal level of privacy or data protection (D-d*) automatically results, and the data disclosed equals d**. At the same time, in the service market x, the optimal quantity x** is supplied and demanded.

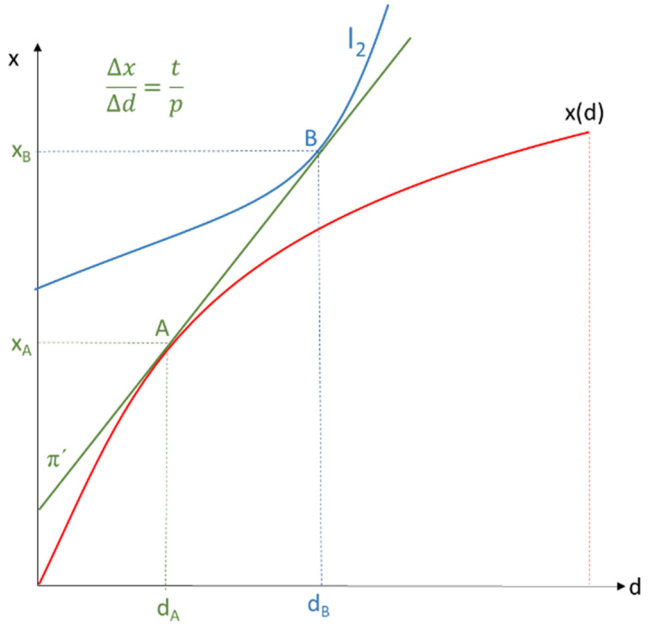

Figure 7. Disequilibrium in the data market.

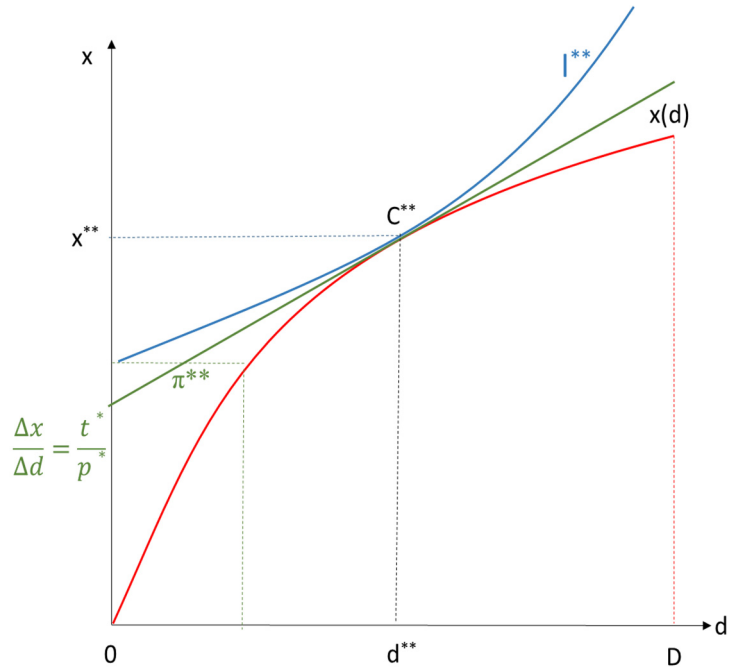

Figure 8. Equilibrium in the data market.

An obligation under the new ePrivacy Regulation to finance the provision of good x by a monetary price in addition to "financing from the disclosure of data" can be explained with the help of Figure 9. The starting point would be D, where all relevant data are disclosed (no data protection), but data-driven goods to the extent of x_D are consumed "free of charge" in return. If, however, the obligation to make a counteroffer with a monetary price p were to take effect, the utility-maximizing point C** would be reached under competition. At the market price for x in the amount of p*, the monetary amount $(x_D\text{-}x^{**})p^*$ would be paid for the smaller but optimal quantity x**; data would only be given up to d**, and the data quantity D-d** would not be disclosed. The new "exchange-money-payment" for x is better for the data subject because a higher indifference curve I** is achieved compared to I'. Admittedly, this increase in utility occurs only if the principles of

voluntary exchange apply. Above all, data subjects must be properly informed and data owners must not have monopolistic influence in the data market.

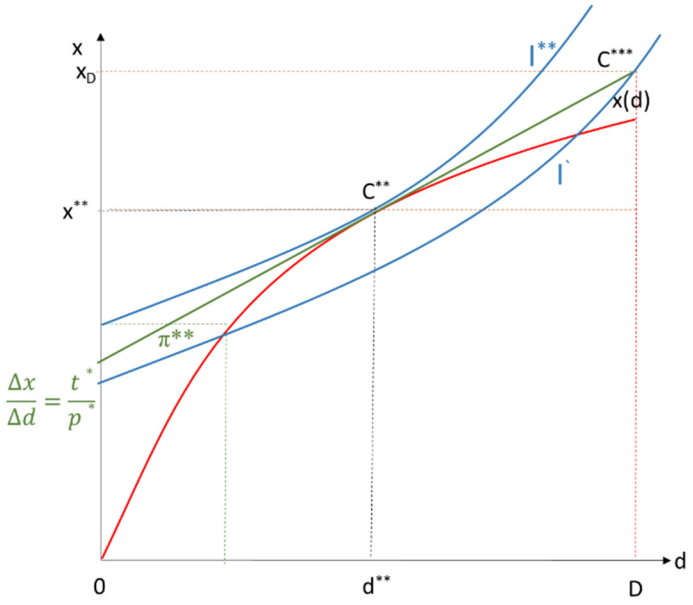

Figure 9. Monetary offer.

This simple welfare economic model of optimal privacy, in which the data owner and data subject appear as one person, shows that there is an optimal level of privacy. This optimal level can be achieved either by paying a monetary price for the Internet service or by giving up data. Admittedly, the conditions of voluntary exchange must be met for this to occur, including well-informed data subjects and competition among data owners.

4. An Experiment with Lueneburg Students

In an online experiment[2] with students from Lueneburg, we tested whether hypothetically offering a monetary price for Internet services instead of having users pay by surrendering their data reduced the extent of cookie fatigue. The control group could only use the artificial intelligence DeepL[3] if they paid with their data or abstained from the service altogether. Treatment Group 1 was additionally given the option to use DeepL without loss of privacy in exchange for a monetary fee, with respondents themselves able to designate their maximum willingness to pay. Treatment Group 2 could instead vote for a "campus" license of DeepL at the annual fee of EUR 10 via a student poll.

In June and July 2021, students were asked about data privacy in an online experiment in which they could hypothetically decide whether to use the Artificial Translation Intelligence DeepL. Randomly, they also had the option of paying a monetary fee instead of revealing their data. DeepL offers a free version that only allows a few texts to be translated and where the translated texts are stored for up to eight years. The latter is done to improve the software; however, there are also various service packages offered on a subscription basis for a fee where texts are deleted immediately after translation.

All students at Leuphana University Lueneburg were invited to participate in the online experiment, which they were informed of via the university's central mailing list. In addition, information about the experiment was published in the university's weekly general information medium, Leuphana Facetten. The General Students Committee (ASTA) reported on the experiment in its newsletter as well as issuing a call for students to participate. In some economic courses conducted via Zoom meeting, students were reminded of the survey or asked to participate via email. A total of 190 students opened the link to the survey, and 147 of them answered.

In the online experiment, participants could choose between different options in the first question, depending on which variant they were randomly assigned to:

- In the Control Group, if they disclosed their data via cookies, they could translate up to three Word or PowerPoint documents with a maximum file size of 5 MB. They could also build a glossary with up to 10 entries. DeepL is subject to the European Data Protection Regulation (DSGV0), so its servers are subject to European data protection law. DeepL requires consent from its users via cookies to store submitted and translated texts for up to 8 years. The purpose of the data storage is to improve the translation performance of DeepL artificial intelligence. If the user refused consent via cookies, they were not able to use DeepL. As another option, Treatment Group 1 respondents could indicate their maximum willingness to pay for the "starter" subscription with extended data protection. This option included unlimited translation opportunities of texts (by inserting it into the mask) and for five documents (Word or PowerPoint) per month. They could choose between a maximum annual fee of EUR 2, EUR 4, EUR 7, EUR 10, EUR 13, EUR 16, EUR 20, EUR 50, or EUR 72 to receive access. Treatment Group 2 had the option to vote for collective student access (campus license). Collective student access could be obtained if DeepL, or a comparably powerful translation software, was included in the next student ballot vote. If the majority of students voted for the software, all students would have access to the software via computers at the university or a VPN. After consultation with DeepL, access would be available at an annual student price of about EUR 10. If a majority of students preferred this option, all students would have to pay the additional fee. The possibility of providing access to the DeepL software in this manner has already been discussed with individual representatives of ASTA.

Figure 10 graphically depicts the decision situations for the control group and the two treatment groups. At the beginning of the experiment, participants were randomly assigned to one of the three groups. Subsequently, all three groups took part in an identical control survey.[4] Data collected included personal information: age, gender, study phase and course, semester, occupation, and financial situation. Participants were also asked about the relevance of language skills and their existing knowledge and preferences of data privacy. They answered questions on their use of fee-based, privacy-friendly services such as search engines, messenger services, and e-mail programs, as well as about current data privacy policy issues.

All participants who expressed a sufficient willingness to pay at least the EUR 10 annual fee in Treatment Group 1, or who voted for the campus license in Treatment Group 2, were given the chance to subscribe to a DeepL service package with good data protection for one year at a very low price. Any interested party would have had to pay EUR 10 out of their pocket, and the difference of the actual annual amount of just under EUR 72 would then have been paid by university budget funds. Since the potential grant requirement from this experiment was over EUR 1000, a lottery was drawn.[5]

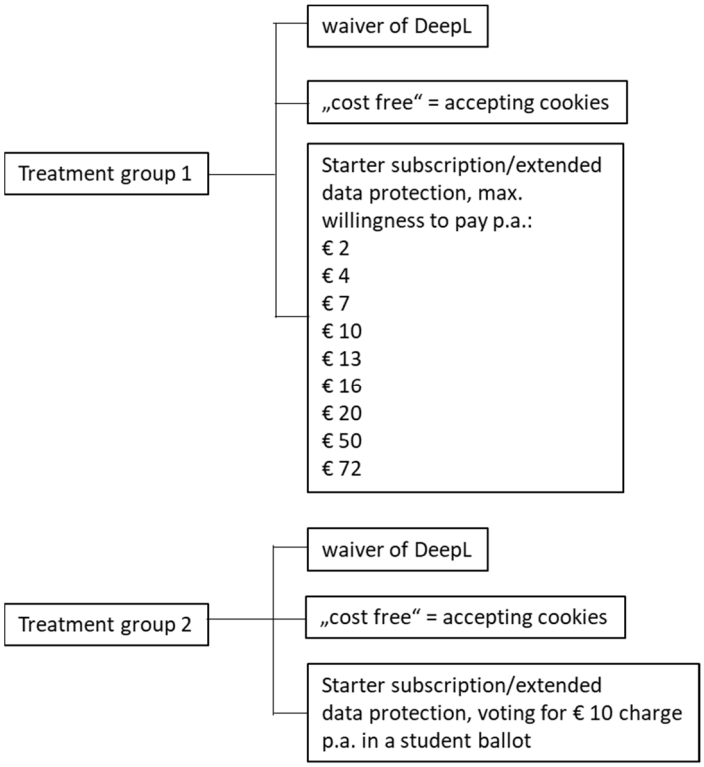

Figure 10. Experimental design.

As the online experiment used a between-subjects design, the results can be interpreted as follows, as shown in Figure 11:

- In the Control Group, potentially 79.31 percent of participants might be affected by cookie consent fatigue because they accept the violation of their privacy by accepting cookies perhaps without really wanting them. In Treatment Group 1, which worked with hypothetical willingness to pay options, the cookie "share" dropped to 30.43 percent. This effect was even stronger in Treatment Group 2 (27.91 percent), in which a price was specified that must be accepted for a transaction, as it is in markets. Admittedly, considering Treatment Group 2, only a collective demand was simulated because it would require a majority vote for sufficient willingness to pay before it would become effective. Comparing the two treatment groups with the control group, the willingness to accept cookies and thus to pay for the service by surrendering personal information decreases to a considerable extent. In the Control Group, 20.69 percent refrained from using DeepL because their privacy concerns were too great. The implicit price was too high compared to the benefits of DeepL. By expressing a positive willingness to pay, DeepL can be used without the loss of privacy (Treatment Group 1), wherein 4.35 percent of the group directly refrained from using the software. With the option of using DeepL as a campus license in the amount of EUR 10 per year (Treatment Group 2), only 2.33% of this group still wanted to abstain from DeepL completely. Compared to the Control Group, both treatment groups favored having a monetary choice to disclose or not disclose data. Microeconomically, this is not a surprising result, since an expansion of the possibility set enables utility increases, with the exception of the rare case of a corner solution. We must also note that the starter subscription also promises a higher level of service for translations compared to the "free" option. This "confounding variable" was unavoidable to improve the external validity of the experiment.
- If we compare the two treatment groups, we see that almost equal proportions opted for the third alternative: 65.21% in Treatment Group 1 with a positive willingness to pay and 69.77% with the acceptance of DeepL at an annual price of EUR 10 in Treatment

Group 2. This parallel evaluation is the other side of the coin to the reduction in the acceptance of cookies or the waiver of DeepL in the control group. However, in Treatment Group 1, 34.78% expressed a willingness to pay less than EUR 10. However, the starter version of DeepL would only be realistically available to students at the "market price" of EUR 10. The first interpretation of this result would be that the true willingness to pay of these 16 persons is not sufficient to pay a realistic monetary price in order to use DeepL without giving up privacy; the only option would be to accept cookies or to do without the service. The second interpretation would be that the participants in Treatment Group 1 did not know what a fair market price for this service would be; if there had been a price signal of EUR 10, they would also have revealed this as a willingness to pay. The second interpretation supports the assessment that the introduction of a market price as a third alternative helps subjects to implement their true preference for privacy.

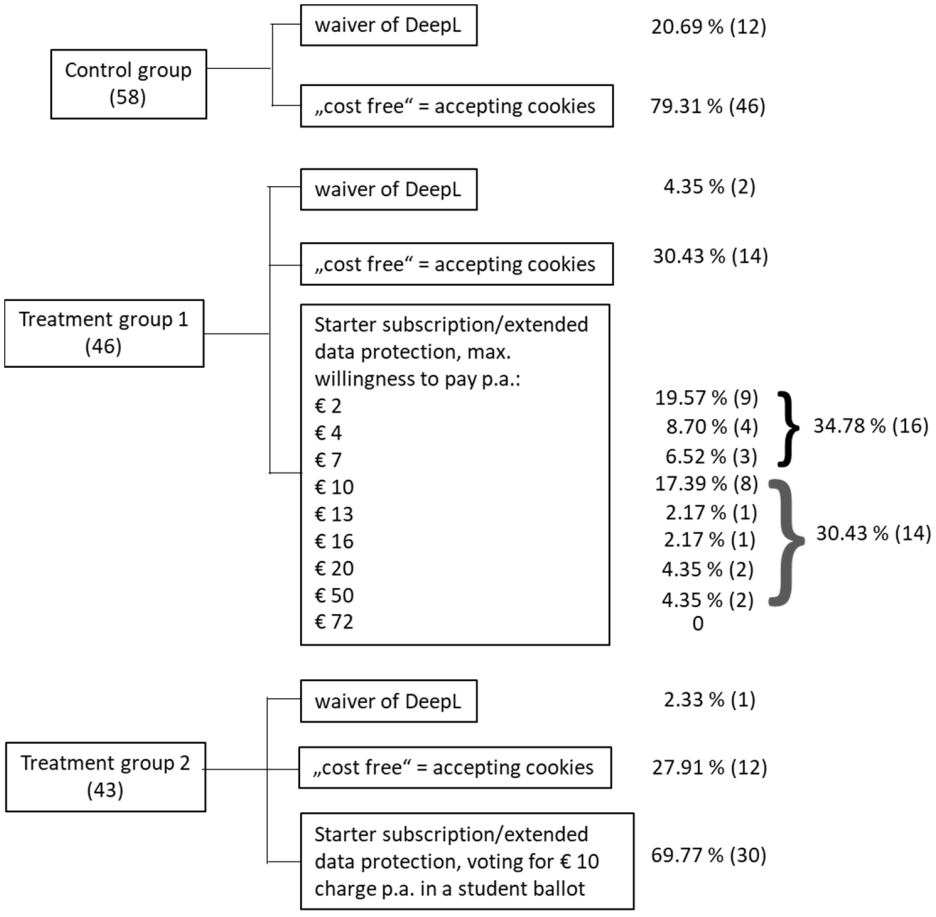

Figure 11. Experimental results.

The introduction of a monetary price for more data protection led to a decrease in the acceptance of cookies or the complete abandonment of the Internet service DeepL in both treatment groups; broader options for action for the participants benefitted them.

A total of 126 participants provided information on their student phase: 62.7 percent were in the bachelor's program, 26.98 percent were in the master's phase, 7.94 percent were doctoral students, and 2.38 percent were enrolled in the Professional School. According to their disclosed major subject, business economists with BWL/IBAE majors represented the largest groups with 13 participants, followed by 11 environmental scientists and economists whose majors included Environmental Sciences, Global Environmental, and Sustainability Studies. At the master's level, 10 respondents came from management programs, and 6 came from Public Economics, Law, and Politics (PELP). Otherwise, there were a few

clusters from other majors, either identifying bachelor minors, doctoral programs, or the Professional School.

On average, respondents (Table 1) were slightly older than 25, in their fifth semester of study, worked slightly less than two days in addition to their studies, and were relatively satisfied with their student living situation—just under 7 out of 10. Of the 124 respondents, 109 grew up with German as their native language; 2 had English as their native language; and 1 each came from Belgium, Cambodia, Venezuela, Ghana, Italy, Russia, and Syria.

Table 1. Socio-demographic characteristics.

	Mean	Standard Deviation	Min.	Max.	Observations
Subject semesters	4.47	2.93	1	20	115
Age of participants	25.42	5.23	18	45	123
Average weekly working time in hours outside of studies	14.87	9.13	2	40	93
Overall life satisfaction (not satisfied at all = 0, totally satisfied = 10)	6.85	1.67	0	10	120

Own calculations with Stata 16.0.

On average, participants attended just over 9 years of school with English classes, and 22 were in an English-language school. According to the standard school grading scale[6], 118 respondents rated their overall English proficiency at an average of 1.98. They considered their reading skills to be at a 1.71 and their writing skills to be at a 2.35. Of the 85 respondents who could reveal English placement according to the European Framework of Reference, 3 indicated a B1 level, 22 indicated a B2 level, 51 indicated a C1 level, and 19 indicated a C2 level. The general financial situation of 120 respondents looked quite positive: 22.5 percent did not have to worry about their finances; 41.67 percent could save a little; and 30 percent could just about manage with their budget. Only seven participants, 5.83 percent, never had enough money.

The relevance of English skills in studies and in private life is shown in Table 2. Foreign languages other than English played almost no role. In study, English skills were frequently or even daily relevant for two-thirds of the participants. The need to read English-language texts was particularly strong, while the need to write a thesis in English was relatively less important. We used the weights 4 for "daily", 3 for "frequently", . . . , and 0 for "never", In summing up these values over all study-relevant needs, the result was a mean value of 13.14, i.e., an average of a 2.6 per need group. In the private sector, "often" was named the most frequently, and "rarely" second most frequently. The private relevance for English language skills thus seems lower than the student relevance.

The questionnaire also recorded preferences on data privacy; on legal knowledge of data privacy; on the actual use of fee-based, data privacy-friendly services; and on views of current data privacy policy issues. Preferences were indicated by putting basic rights guaranteed in Germany into a preference order, within certain limits:

- freedom to organize in a trade union;
- freedom of religion;
- right to informational self-determination;
- freedom to choose an occupation;
- protection of property;
- freedom of demonstration.

Table 2. Relevance of English skills.

	Never	Rarely	Regularly	Often	Daily	No Answer
	in Percent					
In my studies, I need English . . .						
To read the relevant literature (textbooks, journal articles, etc.)	0.83	-	8.33	25.00	65.83	-
In the course because English is the language of instruction	4.17	15.00	15.83	25.83	38.33	0.83
For written or oral exams and tests	8.33	19.17	20.83	21.67	21.67	3.33
For English-language seminar papers or term papers	8.33	18.33	14.17	25.83	31.67	1.67
For English-language theses (bachelor's thesis/master's thesis/dissertation/paper)	20.00	10.83	8.33	10.00	35.00	15.83
In private settings (travel, families, friends) do I need to know English?	0.83	19.17	27.50	35.83	16.67	-

A total of 120 answers. Own calculations with Stata 16.0.

Of 115 respondents, 40 ranked the right to informational self-determination first, and an additional 23 ranked it second. The majority thus put the fundamental right to informational self-determination, which underlies data protection, very near the front of their preferences.

To generate a measure of preference order, if the right to informational self-determination was ranked first, it was weighted by a factor of 1. If it was ranked second, it was weighted by a factor of 0.8. If it was ranked third, it was weighted by a factor of 0.6. If it was ranked sixth or last, it was weighted by a factor of 0. On average, a ranking of 0.7 was obtained with a standard deviation of 0.29 out of 115 observations.

Six legal statements were introduced in the questionnaire, and respondents were asked to check for correctness. Statements 3 and 5 were incorrect, and the remaining four were correct. If the number of correct statements was summed up, an indicator of the knowledge of data protection law was obtained. On average, a knowledge indicator of 3.26 was obtained, with a standard deviation of 1.1, minimum of 0, and maximum of 6. Anyone with data privacy concerns can already use paid services, which are much better at respecting the privacy of users than the "free" services are. A total of 4 participants used paid search engines, 21 used paid messenger services, and 32 used paid e-mail management programs.

Table 3 shows the participants' views on current privacy policy issues. In line with the high preference for the fundamental right of informational self-determination, all statements advocating high data protection were rated as very good or at least useful by many participants. It is interesting to note that the two conceivable innovations of ePrivacy, default setting via browser and obligation to provide paid alternatives when cookies are used, were still met with approval, but the latter to a significantly lesser extent.

Another question was whether respondents who can be shown to have consciously chosen DeepL differed from those who rejected DeepL when controlling for group characteristics. Table 4 shows this for feasible mean comparisons. Only the following mean differences were significant:

- Older students were more likely to choose DeepL than younger students.
- Increasing relevance of English skills for theses and in private life increased the willingness to take up a DeepL offer.
- Better English language writing skills increased the demand for DeepL.
- Higher perceived relevance of the fundamental right to informational self-determination also contributed to more DeepL.

Table 3. Opinions on data protection policy.

	Unnecessary	More Costs Than Benefits	Neutral	Useful	Very Good	Observations
	In Percent					
Data of European users may only be used in the USA according to European standards	0	7.27	9.09	31.82	51.82	110
Infection control software, such as the Corona-Warn app, may only be approved under valid, strict data protection	3.60	14.41	16.22	18.92	46.85	111
Consent to the use of cookies may only be obtained via an opt-in option (the user must explicitly allow data use)	2.70	11.71	11.71	30.63	43.24	111
Individual privacy preferences should generally be able to be set via the browser; the case-by-case consent to data use would then no longer apply	1.82	4.55	15.45	33.64	44.55	110
In addition to the free use of Internet services through the disclosure of data, every user should also be given the opportunity to use the services by paying a fee instead of disclosing data.	3.77	7.55	25.47	29.25	33.96	106

Own calculations with Stata 16.0.

Cross-tabulations with various dummy variables such as gender, native language English, correct legal knowledge, or use of various paid services did not lead to any significant correlations, so their reproduction was omitted.

Table 5 describes the partially significant results of the logit estimates to explain why experiment participants chose to accept cookies. Additional non-significant results can be found in Appendix B. According to Model 1 of Table 5, the probability of accepting cookies with an average marginal effect decreased by slightly more than 50 percentage points when a decision was made in Treatment Group 1 or Treatment Group 2 compared to the Control Group. These highly significant probabilities even increased to about 75 percentage points when controlling for participant-specific effects (Model 3 and 4). In most cases, these effects were completely insignificant (see Appendix B) or were nonrandom only for some variables in Models 2–4 (Table 5). As the need to read English-language texts during study increased, the probability of accepting a cookie decreased: a participant who needs to read English-language literature regularly instead of infrequently, or often instead of regularly, was ceteris paribus 34 percentage points (pp) less likely to choose cookies; however, this relationship was only different from zero at the 10 percent confidence level. If the relevance of English-language exam units increased by one unit, at the 5 percent significance level, the probability of acceptance increased by 25 pp. From an economic theory perspective, it is difficult to find an explanation for the opposing effects. Model 3 suggests that one more year of English language schooling leads to a small average marginal effect of 5 pp (error probability below 5 percent). With respect to current discussions on data privacy policy, according to Model 4, there was a one unit higher endorsement: (a) that data of European users in the USA may only be used according to European standards would increase cookie acceptance by 10 percentage points; (b) that cookie consent may only be obtained as opt-in options would decrease cookie acceptance by 12 pp; and (c) that a service must be offered for a fee in addition to fee-free data use would decrease cookie acceptance by a very small 2 percentage points; however, these effects are at best only at the 5 percent level different from zero. Essentially, however, it must be noted that cookie acceptance is driven down by the two treatments. Overall, confounding factors seemed to play a minor role.

Table 4. Mean comparisons.

		Decision		T-Values	Observations For/Against
		For DeepL	Against DeepL		
Age?		26.32	23.13	−2.45 ***	37/24
Subject semester?		4.68	4.25	−0.47	34/24
Working hours per week?		15.4	15.4	0.062	30/18
Years of Schooling in English?		8.89	9.42	0.725	35/24
Sufficient budget?		2.86	2.82	−0.180	22/36
Overall student satisfaction?		6.97	7.00	0.057	37/23
Relevance of English	Reading literature	4.77	4.58	−1.339	35/24
	During courses	3.80	3.9	0.230	
	Exams	3.20	3.30	0.253	
	Writing term papers	3.57	3.71	0.658	
	Writing theses	2.74	2.66	−0.144 **	
	Private settings	0.23	0.08	−1.462 **	
Years of schooling in English		8.89	9.42	0.725	35/24
English proficiency in school grades	Overall	2.00	1.65	−1.510	35/23
	Reading	1.69	1.52	−0.770	
	Writing	2.4	1.9	−2.071 **	
Relevance of informational self determination		0.76	0.61	−2.010 **	35/24
Correct answers on data protection law		3.47	3.30	−0.574	34/23
Data protection policy	European data standards in the USA	4.85	4.73	−0.313	34/22
	Infection control software only under strict data protection	4.24	4.59	0.781	34/22
	Consent to cookies only as an opt-in option	4.32	5.09	1.972 *	34/22
	Privacy preferences via browser	4.97	4.45	−1.284	34/22
	Additional obligation to offer against payment	4.09	4.32	0.545	33/22

* $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$; own calculations with Stata 16.0.

The estimates of Table 5 still need to be evaluated for their model goodness. To measure their explanatory power, one can ask what percentage of the observations would have been correctly classified (cookies predicted to be rejected/actually rejected or cookies predicted to be accepted/actually accepted) by picking observations quasi-randomly compared to the prediction of the model. In Model 1, the proportion of those correctly predicted increased from 48.9 to 74.2%, for Model 2 from 51.5 to 79.8%, for Model 3 from 50.8 to 80 %, and for Model 4 from 47.1 to 77.9%; in total, a not insignificant explanatory power in each case. Furthermore, as a rule of thumb, the appropriateness of the logit model can be measured by relating the number of cases for the lower value of the variable to be explained, in this case, acceptance of cookies versus rejection, to the number of explanatory variables (without the constant) and trusting in the appropriateness if a value of 10 or greater is given (Van Smeden et al. 2016). Models 1 and 3 reached this threshold, whereas models 2 and 4 did not. In Model 3, there was also evidence of collinearity, as the covariance value between schooling in an English-speaking country and studying at an English-speaking

university was 0.16. For all other variables, the covariance values remained well below 0.1, so collinearity probably did not play a role there.

Table 5. Logit-regression cookie decisions.

		Model 1	Model 2	Model 3	Model 4
Treatment Group 1 = 1		−0.54 *** (0.05)	−0.67 *** (0.04)	−0.72 *** (0.03)	−0.73 *** (0.03)
Treatment Group 2 = 1		−0.57 *** (0.05)	−0.83 *** (0.03)	−0.76 *** (0.03)	−0.82 *** (0.02)
Relevance of English	Reading literature		−0.34 * (0.19)		
	During courses		−0.09 (0.27)		
	Exams		0.25 ** (1.37)		
	Writing term papers		−0.15 * (0.20)		
	Writing theses		0.09 (0.37)		
Years of schooling in English				0.045 ** (0.09)	
Stay in English-speaking school?				−0.21 (0.32)	
Study at English-speaking university?				0.008 (0.49)	
Data protection policy	Data of Europeans in the USA only according to European standards				0.10 * (0.31)
	Infection control software, such as the Corona-Warn app, may only be approved under valid, strict data protection				0.26 (0.16)
	Consent to cookies only as an opt-in option				−0.12 ** (0.11)
	Privacy preferences via browser, case-by-case consent not required				−0.07 (0.14)
	In addition to free use, obligation to offer against payment				−0.02 * (0.16)
Constant–odds ratios		3.83 *** (1.25)	371.33 *** (756.13)	1.29 (0.786)	25.72 ** (42.48)
Wald test (X^2, (p-values))		30.99 *** (0.000)	23.56 *** (0.001)	30.22 *** (0.000)	33.50 *** (0.00)
Correctly classified with constant/with full model in percent		48.9/74.2	51.5/79.8	50.4/80.0	47.1/77.9
Events per variable (EPV)		31	6.9	11.4	7
Observations		147	99	115	104

Dependent variable: cookie accepted = 1, otherwise = 0. Values marginal average effects, robust standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$. Own calculations with Stata 16.0. English native language omitted.

5. Summary and Conclusions

When using Internet services, private households often feel compelled to agree to the use of cookies in order to be able to use the services offered for "free". This is known as "cookie consent fatigue". They pay for services via the implicit price of providing data

about themselves. They may not even want these data-based transactions, or they may prefer to be charged a monetary fee.

A legislative process is currently ongoing in the European Union to add a new privacy regulation to the 2018 General Data Protection Regulation. One innovation of the new regulation would be that service providers on the Internet, who currently must obtain the consent of their users via an opt-out provision, must always provide a fee-based alternative without disclosing data.

A simple economic exchange model shows that users, as data subjects, are basically faced with the choice of paying a monetary price for the service and preserving their privacy, or putting privacy preferences second and using Internet services "for free" while accepting a loss of utility in the form of reduced privacy. If we assume that a person is a data user and thus provides an Internet service, and also acts as a data subject who perceives the surrender of privacy as a loss of utility, a market equilibrium is reached with a utility-maximizing demand for data privacy and an optimum supply of Internet services for which a charge is made. The individual demand for data privacy only coincides with the socially optimal demand if there is competition in the markets for data and Internet services, or if users are sufficiently informed. If the current data privacy policy does not sufficiently protect the privacy of users, an obligation to offer services for a fee can make users better off than the current legal system.

In an online laboratory experiment with students at Leuphana University, the focus was on data privacy while using the artificial intelligence DeepL, which can translate foreign-language texts into many languages. Today, DeepL offers several usage options. One option is to use the translation function to a limited extent free of charge, but users "pay" for usage by agreeing to DeepL storing and reusing the texts entered for up to eight years. If you refuse this consent, free use is not possible. DeepL also offers paid subscriptions, where users' texts are immediately deleted.

All students at Leuphana University in Lueneburg were invited to participate in the online experiment. A total of 190 students opened the link to the survey and 147 of them answered. The Control Group received exactly the same conditions as DeepL without payment: exposure of data for eight years with limited possibility of translation or no translation service if users "insisted" on the privacy of their data. In Treatment Group 1, respondents could indicate their maximum willingness to pay for the so-called starter subscription with extended data protection as an additional option by choosing between different maximum annual fees. In the Treatment Group 2, there was the other option of voting for collective student access to the starter subscription of DeepL at an annual price of EUR 10 (campus license), with access available via university computer or VPN. If this option was accepted by the majority of students in a student vote, all students would have to pay the additional fee. To motivate the participants realistically, they could get the starter subscription available on the market and had to pay only EUR 10 per year with the university supplementing the remainder of the actual costs.

In this between-subjects design, cookie consent fatigue may have affected almost 80 percent of participants in the Control Group, while in Treatment Group 1, the cookie "share" dropped to just over 30 percent, and even more so in Treatment Group 2, by about 28 percent. Compared to the Control Group, both Treatment Groups showed that the option of not only being able to choose between not using cookies or using them with a low level of data protection, but also being able to pay directly for the service, allows users' preferences to take better effect. The descriptive values, the bivariate tests, and the limited feasible logit estimates suggest that the decrease in cookie consent fatigue cannot be explained by a different composition of the three subgroups but are "caused" solely by the interventions of the treatment.

Limits to the validity of the experiment are, of course, that students certainly deal differently with the Internet and with data protection than the general population, and that translation software is likely to be of greater importance to them. It also remains to be considered that the starter subscription has a higher scope of services for translations

compared to the "free" scheme. With the student vote, one makes the demand for DeepL a collective good, although translation aids are certainly a private good; however, it is only through collective provision that one arrives at annual fees that are potentially accepted by students.

In summary, in the theoretical model under competition the optimal demand for privacy would be realized. With the obligation of ePrivacy regulation to declare prices for allowing privacy, optimal privacy demand is made possible. Compared to the status quo of presumed cookie consent fatigue, there is an improvement in the sense that users receive more choices. The experiment with Leuphana students shows this very clearly; however, the experiment also shows that the price for privacy is the decisive parameter. If the service providers charge access fees at an astronomical price, the regulatory mandate of choice does not help much. Only if there is sufficient competition in markets for data and for online services will there be an improvement.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be requested from the author by e-mail.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Optimal Data Protection

Let the production function of the data owner be

$$x = Ad^{\beta}. \tag{A1}$$

The utility function of the data subjects can be written as

$$U(x, D-d) = x^{\alpha}(D-d)^{(1-\alpha)}. \tag{A2}$$

This results in the optimization problem:

$$\max_{x,d} x^{\alpha}(D-d)^{(1-\alpha)} \text{s.t.} x = Ad^{\beta}, \tag{A3}$$

which leads to the Lagrange function:

$$\mathcal{L} = x^{\alpha}(D-d)^{(1-\alpha)} + \lambda(x - Ad^{\beta}) \tag{A4}$$

After forming the first-order conditions and algebraic rearrangements, we obtain

$$d^{*} = \frac{\alpha\beta D}{1-\alpha(1-\beta)} \qquad x^{*} = A\left(\frac{\alpha\beta D}{1-\alpha(1-\beta)}\right)^{\beta}. \tag{A5}$$

Decisions of W as Data Owner and Data Subject

Data owners as profit-maximizing producers optimize:

$$\max_{x,d} px - td \qquad \text{s.t.} x = Ad^{\beta}, \tag{A6}$$

Inserting the constraint yields

$$\max_{d} pAd^{\beta} - td. \tag{A7}$$

This determines the optimal data demand

$$d^{**}_{\text{Demand}}(t,p) = \left(\frac{\beta pA}{t}\right)^{\frac{1}{(1-\beta)}}, \tag{A8}$$

and this result inserted into Equation (A1) leads to the profit-maximizing supply of Internet services:

$$x^{**}_{\text{Supply}}(t,p) = A\left(\frac{\beta pA}{t}\right)^{\frac{\beta}{(1-\beta)}}. \tag{A9}$$

It arises as profits of the data owner:

$$\pi(t,p) = px_{\text{Supply}} - td_{\text{Demand}}. \tag{A10}$$

Equations (A8) and (A9) inserted into (A10) as well as transformed leads to the equation

$$\pi(t,p) = (1-\beta)(Ap)^{\frac{1}{(1-\beta)}}\left(\frac{\beta}{t}\right)^{\frac{1}{(1-\beta)}}. \tag{A11}$$

As a data subject, W earns income I from profit as a data owner and from giving up data,

$$I = td + \pi(t,p). \tag{A12}$$

which is remunerated at t per unit.

Assuming that at most the data set D can be "used" by the data subject, her utility maximization problem arises:

$$\max_{x,d} U(x, D-d) = x^{\alpha}(D-d)^{(1-\alpha)} \text{s.t.} px = td + \pi(t,p). \tag{A13}$$

From the application of the usual Lagrange maximization follows the utility maximizing data set:

$$d^{**}_{\text{Supply}}(t,p) = \alpha D - \frac{(1-\alpha)\pi(t,p)}{t} = \alpha D - \frac{(1-\alpha)(1-\beta)}{\beta}\left(\frac{\beta pA}{t}\right)^{\frac{1}{(1-\beta)}} \tag{A14}$$

and as utility-maximizing demand for Internet services:

$$x^{**}_{\text{Demand}}(t,p) = \frac{\alpha}{p}(tD + \pi(t,p)) = \frac{\alpha t}{p}\left(D + \frac{(1-\beta)}{\beta}\left(\frac{\beta pA}{t}\right)^{\frac{1}{(1-\beta)}}\right). \tag{A15}$$

Equilibria in data and Internet markets prevail if Equations (A8) = (A14), and (A9) = (A15), respectively, are satisfied; the first condition ensures market clearing in the data market, and the second in the market for Internet services. Calculating the market clearing in the data market, the equilibrium data price is as follows:

$$t^{**} = \beta A\left(\frac{1-\alpha(1-\beta)}{\alpha\beta D}\right)^{(1-\beta)} p. \tag{A16}$$

The same is true if we calculate the market clearing in the Internet market.

Solving Equation (A16) for p and substituting this into the profit-maximizing supply for Internet services (Equation (A9)) (alternatively into the optimal demand) yields

$$d^{**} = \frac{\alpha\beta D}{1-\alpha(1-\beta)}, \tag{A17}$$

which is identical to the welfare-optimal quantity in the data market d* from Equation (A5).

Appendix B

Table A1. Logit-regressions cookie decisions—non-significant confounders.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group 1 = 1	0.07 *** (0.04)	0.03 *** (0.02)	0.07 *** (0.04)	0.05 *** (0.03)	0.06 *** (0.03)	0.08 ** (0.04)	0.08 *** (0.04)	0.07 *** (0.04)
Treatment Group 2 = 1	0.06 *** (0.03)	0.06 *** (0.04)	0.07 *** (0.04)	0.05 *** (0.03)	0.05 *** (0.03)	0.08 *** (0.05)	0.07 *** (0.04)	0.07 *** (0.04)
Master program, yes?	1.02 (0.51)							
PhD-candidate, yes?	0.62 (0.43)							
Professional school, yes?	1.47 (1.95)							
Age?		1.03 (0.07)						
Women yes, men and divers no		0.81 (0.50)						
Subject semester?		1.00 (0.08)						
Working hours per week?		0.97 (0.03)						
Relevance of budget constraint?		0.58 (0.24)						
Overall life satisfaction?		0.95 (0.15)						
Sum of English relevance?			0.98 (0.04)					
English proficiency in school grades: overall				0.94 (0.48)				
English proficiency in school grades: reading				1.31 (0.54)				
English proficiency in school grades: writing				1.37 (0.50)				
Common European Framework of Reference for Languages (CEFRL)					0.92 (0.90)			
Sum of the correct answers on data protection law						0.78 (0.15)		
The basic right to informational self-determination–Aggregate value							1.04 (0.79)	
Use of fee-based search engines?								0.60 (1.22)

Table A1. *Cont.*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Use of fee-based messenger services?								0.93 (0.62)
Use of fee-based e-mail programs?								0.40 (0.25)
Constant	5.78 *** (2.59)	39.06 * (84.50)	7.00 *** (4.87)	2.31 (1.67)	8.10 *** (3.71)	10.24 *** (7.29)	4.61 ** (3.17)	6.73 *** (2.90)
Wald test (X^2, (*p*-values))	32.95 *** (0.000)	26.91 *** (0.001)	31.25 *** (0.000)	41.53 *** (0.00)	38.34 *** (0.000)	29.76 *** (0.00)	29.68 *** (0.000)	30.93 *** (0.00)
Observations	126	84	120	118	115	112	115	115

Dependent variable: cookie accepted = 1, otherwise = 0. Values odds ratios, robust standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$. Own calculations with Stata 16.0.

Notes

[1] Especially in the long run, it is plausible that more data will lead to increasing marginal returns. In extreme cases, only one data owner will "survive" and a monopoly will emerge. This model variant would lead to the regulation of a natural monopoly. In this model, the focus is solely on data protection policy under competition.

[2] The LimeSurvey software was used for the online experiment.

[3] DeepL is a translation software that allows simultaneous translation of text sections, Word documents, and PowerPoint files, using text from users. Translations are available from English, French, German, and Spanish, as well as from many less-widely used languages. The extensive glossary function allows for individual improvement of one's own texts. Even simultaneous grammatical optimization is possible. In the experiment, two tutorial videos were created for DeepL that could only be accessed via YouTube with a participant-specific link.

[4] The questionnaire can be requested from the author.

[5] According to the available funding and the conditions mentioned above, grant beneficiaries were randomly selected and announced via the institute's homepage using the name–birthday abbreviation recorded in the survey. Eligible individuals not drawn were notified that they were on a waiting list. Only one person (from the waiting list) wanted to "take" the grant, but has not yet claimed it.

[6] Best grade = 1, . . . , worst grade = 6.

References

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. The Economics of Privacy. *Journal of Economic Literature* 54: 442–92. [CrossRef]

Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. 2015. Privacy and Human Behavior in the Age of Information. *Science* 347: 509–14. [CrossRef]

Acquisti, Alessandro. 2010. *The Economics of Personal Data and the Economics of Privacy*. Pittsburg: Carnegie Mellon University.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz. 2020. *The Economic Consequences of Data Privacy Regulation: Empirical Evidence from GDPR*. Cambridge: National Bureau of Economic Research (NBER). Available online: https://www.nber.org/papers/w26900 (accessed on 21 September 2022).

Barth, Susanne, and Menno D. T. De Jong. 2017. The Privacy Paradox–Investigating Discrepancies Between Expressed Privacy Concerns and Actual Online Behavior—A Systematic Literature Review. *Telematics and Informatics* 34: 1038–58. [CrossRef]

Buchner, Benedikt. 2020. Grundsätze des Datenschutzrechts. In *Einführung in das Datenschutzrecht*. 7., überarbeitete und aktualisierte Auflage. Edited by Marie-Theres Tinnefeld, Benedikt Buchner, Thomas Petri and Hans-Joachim Hof. Berlin: De Gruyter Oldenbourg, pp. 220–332.

Burgess, Matt. 2018. The Tyranny of GDPR Popups and the Websites Failing to Adapt. Available online: https://www.wired.co.uk/article/gdpr-cookies-eprivacy-regulation-popups (accessed on 11 July 2022).

Buttarelli, Govanni. 2016. The EU GDPR as a Clarion Call for a new Global Digital Gold Standard. *International Data Privacy Law* 6: 77–78. [CrossRef]

Coase, Ronald H. 1960. The Problem of Social Cost. *Journal of Law and Economics* 3: 1–44. [CrossRef]

Council of the European Union. 2021. Confidentiality of Electronic Communications: Council Agrees Its Position on ePrivacy Rules. Available online: https://www.consilium.europa.eu/en/press/press-releases/2021/02/10/confidentiality-of-electronic-communications-council-agrees-its-position-on-eprivacy-rules/ (accessed on 10 April 2021).

De Hert, Paul, and Vagelis Papakonstantinou. 2016. The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals? *Computer Law & Security Review* 32: 179–94. [CrossRef]

Goldberg, Samuel, Garrett Johnson, and Scott Shriver. 2019. Regulating Privacy Online: The Early Impact of the GDPR on European Web Traffic & E-Commerce Outcomes. Available online: ttps://scholar.google.de/scholar?hl=de&as_sdt=0%2C5&q=goldberg+johnson+shriver&btnG= (accessed on 28 November 2022).

Goldberg, Samuel, Garrett Johnson, and Scott Shriver. 2021. Regulating Privacy Online: An Economic Evaluation of the GDPR. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3421731 (accessed on 21 September 2022).

Goldfarb, Avi, and Catherine Tucker. 2011. Online display advertising: Targeting and obtrusiveness. *Marketing Science* 30: 389–404. [CrossRef]

González, Elena Gil, Paul De Hert, and Vagelis Papakonstantinou. 2020. *The Proposed ePrivacy Regulation: The Commission's and the Parliament's Draft s at a Crossroads? Data Protection and Privacy*. Data Protection and Democracy. Oxford: Hart Publishing, pp. 267–98.

Hermalin, Benjamin E., and Michael L. Katz. 2006. Privacy, Property Rights and Efficiency: The Economics of Privacy as Secrecy. *Quantitative Marketing and Economics* 4: 209–39. [CrossRef]

Jia, Jian, Ginger Zhe Jin, and Liad Wagman. 2018. *The Short-Run Effects of GDPR on Technology Venture Investment*. Cambridge: National Bureau of Economic Research. Available online: https://www.nber.org/papers/w25248 (accessed on 21 September 2022).

Jia, Jian, Ginger Zhe Jin, and Liad Wagman. 2021. The Short-Run Effects of the General Data Protection Regulation on Technology Venture Investment. *Marketing Science* 40: 661–84. [CrossRef]

Johnson, Garrett, Scott Shriver, and Samuel Goldberg. 2020. Privacy & Market Concentration: Intended & Unintended Consequences of the GDPR. Available online: https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3477686 (accessed on 21 September 2022).

Larouche, Pierre, Martin Peitz, and Nadya Purtova. 2016. Consumer Privacy in Network Industries, Centre of Regulation in Europe (CERRE). January 25. Available online: https://cerre.eu/publications/consumer-privacy-network-industries/ (accessed on 10 April 2021).

Laudon, Kenneth C. 1996. Markets and Privacy. *Communications of the ACM* 39: 92–104. [CrossRef]

Lefrere, Vincent, Logan Warberg, Cristobal Cheyre, Veronica Marotta, and Alessandro Acquisti. 2020. The Impact of the GDPR on Content Providers. *The 2020 Workshop on the Economics of Information Security*. Available online: https://scholar.google.de/scholar?hl=de&as_sdt=0%2C5&q=lefrere+waberg&btnG= (accessed on 21 September 2022).

Machuletz, Dominique, and Rainer Böhme. 2019. Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR. *arXiv* arXiv:1908.10048. [CrossRef]

Nechyba, Thomas J. 2018. *Intermediate Microeconomics: An Intuitive Approach with Calculus*. Andover and Hampshire: Cengage Learning EMEA.

Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer. 2020. European Privacy Law and Global Markets for Data. Center for Law & Economics Working Paper Series, 1. Available online: https://scholar.google.de/scholar?hl=de&as_sdt=0%2C5&q=peukert+bechthold&btnG= (accessed on 21 September 2022).

Posner, Richard A. 1978. The Right of Privacy. *Georgia Law Review* 2: 392–422.

Posner, Richard A. 1981. The Economics of Privacy. *American Economic Review* 71: 405–9.

Solove, Daniel J. 2006. A Taxonomy of Privacy. *University of Pennsylvania Law Review* 154: 477–560. [CrossRef]

Solove, Daniel J. 2021. The myth of the privacy paradox. *George Washington Law Review* 89: 1–52. [CrossRef]

Stigler, George J. 1980. An introduction to privacy in economics and politics. *The Journal of Legal Studies* 9: 623–44. [CrossRef]

Utz, Christine, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. 2019. (Un)informed Consent: Studying GDPR Consent Notices in the Field. In Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, London, UK, November 11–15; pp. 973–90.

Van Smeden, Maarten, Joris AH de Groot, Karel GM Moons, Gary S. Collins, Douglas G. Altman, Marinus JC Eijkemans, and Johannes B. Reitsma. 2016. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Medical Research Methodology* 16: 1–12. [CrossRef] [PubMed]

Varian, Hal R. 2002. Economic aspects of personal privacy. In *Cyber Policy and Economics in an Internet Age*. Boston, MA: Springer, pp. 127–37.