

Nam, Sangjun; Kwon, Youngsun

Conference Paper

Quantifying different psychological costs of user behavioral info for overcoming the 'take-it-or-leave-it' condition

31st European Conference of the International Telecommunications Society (ITS): "Reining in Digital Platforms? Challenging monopolies, promoting competition and developing regulatory regimes", Gothenburg, Sweden, 20th - 21st June 2022

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Nam, Sangjun; Kwon, Youngsun (2022) : Quantifying different psychological costs of user behavioral info for overcoming the 'take-it-or-leave-it' condition, 31st European Conference of the International Telecommunications Society (ITS): "Reining in Digital Platforms? Challenging monopolies, promoting competition and developing regulatory regimes", Gothenburg, Sweden, 20th - 21st June 2022, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/265662>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Quantifying different psychological costs of user behavioral info for overcoming the 'take-it-or-leave-it' condition

Sangjun Nam (snam@etri.re.kr)^{a, b}, Youngsun Kwon (yokwon@kaist.ac.kr)^b

^a Technology Policy Research Division, Electronics and Telecommunications Research Institute,
218 Gajeong-ro, Yuseong-gu, Daejeon, 34129, Republic of Korea

^b School of Business and Technology Management, College of Business, KAIST, 291 Daehak-ro, Yuseong-gu,
Daejeon 34141, Republic of Korea

Abstract

Take-it-or-leave-it, in which users have to provide personal information as required by service providers, has been a dominant form of agreement between online service providers and users. The regulators recently began to prohibit dominant online platforms from collecting personal data based on the 'take-it-or-leave-it' basis because this clause is likely to harm consumer welfare without giving users choices for using the service. In order to improve regulatory efficiency, we need to devise more flexible alternative service provisions balancing privacy concerns and enhanced service based on personal preference. To accomplish this goal, we need to understand the users' attitudes related to personal behavioral data collection for both regulators and online platforms. In this context, we aim to estimate the psychological costs that users bear when they need to exchange personal data for service use. Quantifying the perceived cost of personal data collection with monetary reward was common. However, it is not easy to determine whether the perceived cost is high or not because the monetized value of personal data is not self-evident. To address this issue, we consider attention cost, one of the representative inconvenience costs of using free online services in the analysis. This study collects the data using a conjoint survey and estimates the psychological costs of personal data collection using the mixed logit model and latent-class logit model. Our results show that the respondents' perceived cost for overcoming the 'take-it-or-leave-it' condition is heterogeneous, and only one of four respondent segments (around 30% of respondents) perceived it as significant. Moreover, the results suggest that the perceived risks and benefits of personal data collection affect the psychological cost. It implies that privacy calculus theory can be a meaningful framework for understanding users' attitudes toward behavioral data collection on online platforms.

Keywords: Personal data, Online platform, Privacy calculus theory, Information disclosure, Attention cost

1. Introduction

It is common for online platforms to collect users' behavioral data through terms and conditions in exchange for providing free services. However, the users of dominant online platforms often have no choice but to accept the personal data collection conditions to use the platform services due to the lack of alternatives, even if they are concerned about the information disclosure. In this perspective, personal behavioral data collection in terms of using services without ensuring sufficient notices and user options is one of the crucial issues of online platform regulation. Globally, personal behavioral data collection regulation has intensified and becomes more restrictive. In 2019, German regulator Bundeskartellamt ruled that Facebook abused its dominant position by using terms of service to collect users' behavioral data with the 'take-it-or-leave-it' condition. Bundeskartellamt concluded that it is an exploitative abuse because Facebook's terms of service and the manner and extent to which it collects and uses data violate the General Data Protection Regulation (GDPR) to the detriment of users. The German regulator noted that it is difficult for users to decide autonomously on disclosing their data because Facebook is a dominant company (Bundeskartellamt, 2019). In addition, in January 2021, the German Federal Parliament adopted the 10th amendment of Gesetz gegen Wettbewerbsbeschränkungen (GWB Digitalisation Act), which prohibits the company, declared as a 'Paramount Significance for Competition Across Markets (PSCM),' from "making the use of services conditional on the user agreeing to the processing of data from other services of the undertaking or a third-party provider without giving the user sufficient choice as to whether, how and for what purpose such data are processed (§ 19a, para. 2, sentence 2.4.a)." Under the GWB Digitalisation Act, Bundeskartellamt determined that Alphabet/Google and Meta are PSCM in early 2022, which means they are subject to new rules, including those mentioned earlier. Similarly, in 2020, UK regulator CMA (Competition & Markets Authority) suggested 'The Choice Requirement Remedy (CRR),' requiring platforms to give consumers a choice not to share their data for personalized advertising (CMA, 2020).

As the GWB Digitalisation Act already prohibited platform firms, declared as a PSCM, from collecting personal data from other services of the undertaking or a third-party provider based on a 'take-it-or-leave-it' basis, the PSCM platform firms need to develop more refined user consent rules. In addition, the regulators also need to prepare practical criteria for assessing the legal disputes on obtaining user consent for service use. Observing this changing context, we aim to argue that users' different attitudes, measured by psychological costs, towards personal data trading can be essential practical information assisting the regulators and firms. To accomplish this goal, we find evidence that users' attitudes consenting to platform firms' use of personal data are not homogeneous but heterogeneous. Ensuring sufficient privacy options without any trade-offs could increase the welfare of privacy-sensitive users. However, in the case of users who do not much care to provide personal data for free service, their welfare may decline, resulting in a decrease in the efficiency of data-based business models under stringent regulation. Indeed, Facebook suffered by Apple iOS privacy policy change, so-called App Tracking Transparency (ATT), in April 2021, asking users whether they consent to track their behavioral data in iOS based on opt-out. According to the press release, Facebook argued that ATT would decrease its revenue by around \$10 billion by 2022. It implies that the difficulty of collecting personal data due to highly stringent regulations can decrease both the users' welfare and the service providers' business performance. Thus, in order to provide sufficient options for privacy-sensitive users while avoiding inefficiency in terms of social welfare, both regulators and online platforms need to know how users' psychological cost of consenting to the use of their behavioral data for various business purposes varies with users' characteristics.

Previous literature investigated users' privacy behavior and attitude related to personal data disclosure and personal data collection and found that users' attitude varies by users' privacy perception

and context (Hann et al., 2007; Acquisiti et al., 2013; Kehr et al., 2015; Winegar and Sunstein, 2019). However, there is still a lack of research on the users' different attitudes toward the 'take-it-or-leave-it' condition. Considering the users' privacy behavior is context-dependent, users' attitudes related to the 'take-it-or-leave-it' condition may not be similar to previous findings. In this regard, this study aims to investigate the terms of personal behavioral data collection with a few available choices converted into perceived costs. Quantifying the perceived cost of information disclosure and personal data collection with monetary reward or payment is common. However, it is not easy to determine whether the quantified psychological cost is high or not because the monetized value of data privacy is not self-evident (Winegar and Sunstein, 2019). In other words, it can be said that appropriate criteria for comparison in terms of monetary value are needed to determine whether the perceived cost is high in free online service cases. To address this issue, we consider advertising loads, one of the representative inconvenience costs of using free online services, and ad-free subscription fees together in the empirical analysis.

This study uses the two discrete choice models, the mixed-logit and latent-class logit models, to quantify users' psychological cost of behavioral data collection under the 'take-it-or-leave-it' condition with addressing users' heterogeneous privacy attitude. The two models are similar in that each consumer has a utility function, and their preference over the choice set is heterogeneous, but they differ in assuming the probability distribution types of consumer preference (McFadden, 1973; Train, 2009). Moreover, we aim to investigate individual characteristics that affect the heterogeneity, such as perceived risk and benefit of personal data collection. Users' preference data is collected from a conjoint survey, and this study focuses on YouTube, a representative online platform service, providing both ads-based free service and ad-free subscription service. YouTube, therefore, is an appropriate service for comprehensively measuring the psychological costs to users according to differences in advertising load, ad-free subscription fees, and the terms of personal data collection. We also consider the purpose of personal data collection because it is important for determining the scope of the regulation.

This study is organized as follows: Reviewing previous studies related to the meaning of personal data collection in competition policy focused on the transaction between online platforms and users, and this study also reviews previous studies that investigated the psychological cost of information disclosure, including the consent of personal data collection in Section 2. The research methodology and data are presented in Section 3. Section 4 represents the empirical results, and the conclusions of this study are given in Section 5.

2. Literature Review

2.1. The meaning of personal data collection in competition policy

2.1.1. Online platforms' incentive for more personal data collection

Personal data is a valuable resource in digital markets. OECD Privacy Guidelines defined personal data as "any information relating to an identified or identifiable individual (data subject)," which includes behavior data and location data (OECD, 2013). In general, online platform firms provide free services and collect personal data in accordance with terms and conditions to improve the quality of their services, develop innovative services, and monetize (Lerner, 2014; Colangelo and Maggiolino, 2018). As personal data is a valuable input for online platform firms, it can be said that they have an incentive to collect more personal data. In this view, Evans (2013) and Wu (2018) argued that online platform firms compete to secure users' attention with free services. The more users stay longer on the platform, it can acquire more behavior data from users' activity. In this context, previous literature

viewed collecting personal data from the user as a kind of transaction because online platform firms demand personal data in exchange for providing free services, and users consent under terms and conditions to use these services (Lerner, 2014; Newman, 2016; Gal and Rubinfeld, 2017). So, personal data collection is described as a hidden cost for free online services (Newman, 2016; Gal and Rubinfeld, 2017) or a quality dimension that affects users' choices (Cooper, 2012; Kimmel and Kestenbaum, 2014). Stucke and Ezrachi (2016) noted that the lack of competition in the online platform market can lead to increased cost or decreased quality, and it can also be applied to personal data collection. Stucke (2018) argued that a dominant online platform has the incentive to collect personal data above the competitive level to the extent its business model depends on personal data. In this regard, there are some views that excessive data collection is a type of excessive price or unfair condition (Stucke and Robertson, 2019; Robertson, 2020). In summary, online platform firms have an incentive to collect more personal data because it is valuable input to generate revenue and extend their business depending on personal data. So, in the lack of competition, online platform firms may tend to collect personal data excessively above the competition level, such as collecting personal data on a 'take-it-or-leave-it' basis.

2.1.2. Users' behavior related to personal data collection

The following important question is how online platform firms can collect personal data excessively in gaining users' consent. If users cannot adequately manage their data collection, it can be problematic. So, this problem is related to whether individuals can manage their privacy choice about collecting and tracking behavioral data effectively. 'Privacy paradox' is an important concept for understanding the users' data collection attitude. The privacy paradox refers to the contradictory behavior in which users disclose their privacy or personal data even though they are worried about privacy concerns (Kokolakis, 2017). For most users, collecting and processing their data would imply a greater privacy exposure (Reyna, 2018). One of the theories interpreting the gap between privacy attitude and privacy behavior is the 'privacy calculus theory,' which postulates that individuals calculate between the expected loss of privacy and the potential gain of disclosure (Kokolakis, 2017). Based on privacy calculus theory, previous literature investigated the privacy disclosure behavior in social media (Choi et al., 2018), e-commerce (Dinev and Hart, 2006), and willingness to provide personal data for IoT services (Kim et al., 2019). They found that perceived risks and benefits for privacy disclosure negatively and positively related to privacy disclosure intention. However, some literature pointed out that users' privacy choices can be influenced by behavioral biases such as heuristics, incomplete information, information asymmetries, constructed preferences, free effect, and power of default (Zuiderveen Borgesius, 2014; Kehr et al., 2015; Reyna, 2018; Costa-Cabral and Lynskey, 2017; Acquisiti et al., 2020). Also, users' privacy attitudes can be influenced by the positive interface (Kehr et al., 2015), consumer-brand relationship (Hayes et al., 2021), and the lack of alternatives (Hermes et al., 2018). Zuiderveen Borgesius (2014) noted that these biases could steer privacy choices differently because such choices are context-dependent. In privacy, context-dependent means that an individual's privacy attitude can appear differently depending on the situation (Acquisiti et al., 2015).

In summary, users' attitude toward information disclosure is related to context and perceived risks and benefits which are biased by behavioral bias (Barth and de Jong, 2017). These behavioral biases may lead to market failure because firms can easily persuade users to agree to track and collect their behavioral data using biases. Acquisiti et al. (2020) noted that firms could easily persuade users to consent to personal data collection by utilizing the psychological factors mentioned above. For example, under the lack of competition, switching costs and lock-in effects may bias users' privacy choices, and also firms would be likely to reduce the perceived risks by framing their privacy policies in a user-friendly way to give users the idea that they are under complete control over the personal data submitted

to the service (Zuiderveen Borgesius, 2014; Reyna, 2018). This is why regulation with informed consent is not sufficient alone and needs to focus on how information is provided and how choices are shaped (Reyna, 2018). Zuiderveen Borgesius (2014) also argued that as long as the law allows firms to offer 'take-it-or-leave-it' choices, firms can easily persuade people to agree to the tracking. In this background, it can be explained that the regulation which mandates dominant online platforms to provide sufficient options and information when they obtain consent from users for collecting their behavioral data has been introduced.

2.2. Quantifying the users' psychological costs for personal data collection

A representative way to investigate how much users actually care about data privacy is to ask users about their willingness to pay (WTP) or willingness to acceptance (WTA) of personal data collection (Winegar and Sunstein, 2019). There were many empirical studies investigating the WTP for not providing personal data and/or WTA for providing personal data based on monetary reward or payments (Hann et al., 2007; Acquisiti et al., 2013; Savage and Waldman, 2015; Winegar and Sunstein, 2019; Prince and Wallsten, 2021; Alfnes and Wasenden, 2022). These studies empirically found that users' WTP and/or WTA related to privacy disclosure is heterogeneous and context-dependent. Moreover, previous studies found that preference for privacy disclosure can vary among distinct segments (Hann et al., 2007; Savage and Waldman, 2015).

Considering users' information disclosure attitudes varied by contexts and behavioral biases, a more realistic and specific scenario is needed to quantify the psychological cost of users' behavioral data for overcoming the 'take-it-or-leave-it' choices. However, there is still a lack of empirical studies investigating the users' psychological costs of consenting behavioral data collection under 'take-it-or-leave-it' choices. In this background, Which? (2021) estimated the value of CRR for Google and Facebook based on survey data and showed that the value of CRR is £1.09/month (based on WTP) and £4.03/month (based on WTA) under sufficient information. It provided empirical evidence that ensuring consumer choice related to personal data collection can be of value for users. However, it also has a limitation. Which? (2021) used monetary reward and payment to estimate the value of CRR, which is far from a realistic scenario in the nominally free service case. Also, it is difficult to conclude whether this estimated value is high or low because the monetary value of data privacy is not self-evident, as Winegar and Sunstein (2019) noted. One possible approach for addressing this issue is comparing the behavioral data collection condition with more explicit and similar non-monetary factors such as attention cost. Attention cost is a kind of nuisance cost due to the advertisement, and it is also noted as another hidden cost of free online services (Newman, 2016; Gal and Rubinfeld, 2017; Wu, 2018). Unlike personal data collection, paid services for avoiding advertisements are common such as YouTube Premium, ads-based freemium services, and paid Ad-block programs provided by third parties. In this background, Shon et al. (2021) quantified the inconvenience cost of online video advertisement and empirically found that advertising load changes affect users' service choices. For instance, in this approach, a user may choose between the 'take-it-or-leave-it' choice with fewer advertisement loads and ensure choice related to behavioral data collection with more advertisement loads. If users react to the lack of choice related to behavioral data collection as sensitive as changes in advertising loads, it can be evidence that ensuring users' choice is an important factor for users' choice as similar to attention cost in actual markets.

Moreover, we investigate the effect of perceived risks and benefits of personal data collection on users' psychological cost for overcoming the 'take-it-or-leave-it' condition to explain the heterogeneous privacy preferences. Based on the privacy calculus theory, perceived risk (benefit) affects users' privacy disclosure behavior positively (negatively) and positively (negatively) privacy disclosure decreases (increases) the psychological cost. If the perceived risk and benefit affect users'

psychological costs, it can be said that privacy calculus theory is still a useful theoretical framework for understanding users' heterogeneous privacy attitudes when they face a 'take-it-or-leave-it' condition.

The purpose of behavioral data collection also can be an important issue for determining the scope of regulation. In reality, online platform firms separately describe the purpose of personal data collection, such as developing services and marketing, including targeted advertisement in the terms and conditions. Also, CRR proposed by CMA is only limited to personal data collection for target advertisement. Several studies found that a high level of the personalized message led to an adverse reaction even though they have advantages for a user by providing more relevant information (White et al., 2008; Baek and Morimoto, 2012); and Hayes et al. (2021) found that users' willingness to provide personal data for targeted advertisements is related to perceived benefits and risks based on the privacy calculus theory. Considering that collecting behavior data for marketing purposes is related to the targeted advertisement, including personalized messages, it can be said that collecting behavior data for marketing purposes is perceived as more negative or risky than other purposes. In this regard, the psychological costs of behavioral data collection can be higher for the specific purpose of data collection.

3. Methodology and data

3.1. Methodology

This study used a mixed logit model and latent-class logit model from among discrete choice models (McFadden, 1973; Train, 2009). Unlike commonly used discrete choice models such as logit and probit models, the mixed logit model and latent-class conditional logit model assumed that consumers have heterogeneous preferences. The difference between the mixed logit model and latent-class logit model is the probability distribution types of consumer preferences.

In the mixed logit model, it is assumed that each individual n has a utility function related to alternative j in choice set t . The utility function is expressed as Eq. (1).

$$U_{njt} = V_{njt} + \varepsilon_{njt} \quad (1)$$

where V_{njt} is a deterministic utility that is observed by researchers and ε_{njt} is an unknown part treated as random term that is iid extreme value. Assumed that $V_{njt} = \beta_n x_{njt}$ and β_n follows the distribution $\beta_n \sim f(\beta|\theta)$, the probability that individual n chooses alternative j on the t th choice is expressed as Eq. (2), and the probability that individual n chooses T times is expressed as Eq. (3).

$$L_{njt}(\beta_n) = \frac{e^{\beta_n x_{njt}}}{\sum_{k=1}^J e^{\beta_n x_{nkt}}} \quad (2)$$

$$S_n(\beta_n) = \prod_{t=1}^T \prod_{j=1}^J L_{njt}(\beta_n) \quad (3)$$

However, Eq. (3) is the conditional probability when β_n is known. So, Eq. (3) must be made into the unconditional probability that can be estimated through the integral over all possible variables of β_n . The unconditional probability is expressed as Eq. (4).

$$P_n(\theta) = \int S_n(\beta) f(\beta|\theta) d\beta \quad (4)$$

Log-likelihood function expressed as Eq. (5) cannot be used in the mixed logit model because Eq. (4) cannot be solved analytically. Therefore, it is approximated using a simulation method such as simulated maximum likelihood estimation (SMLE). The simulated log-likelihood is expressed as Eq. (6).

$$LL(\theta) = \sum_{n=1}^N \ln P_n(\theta) \quad (5)$$

$$SLL(\theta) = \sum_{n=1}^N \ln \left\{ \frac{1}{R} \sum_{r=1}^R S_n(\beta^r) \right\} \quad (6)$$

where R is the number of replications and β^r is the r -th draw from $f(\beta|\theta)$. This study used the mixlogit package of STATA (Hole, 2007) to implement the mixed logit model. The WTP of each attribute k is calculated from the estimated individual coefficients of each attribute and price as followed eq (7).

$$E(WTP_k) = -E \left[\frac{\partial U_n}{\partial x_{nk}} \right] = -E \left[\frac{\beta_{nk}}{\beta_{n(price)}} \right] \quad (7)$$

In latent-class logit model, it assumes that distribution $f(\beta)$ is discrete with β taking a finite set of distinct value. This is useful if there are M segments in the population, each having its own choice behavior or preferences. The share of the population in segment m is s_m , which the researcher can estimate within the model along with the β 's for each segment. The probability that individual n in segment m chooses is expressed as Eq. (8).

$$S_n(\beta_m) = \prod_{t=1}^T \prod_{j=1}^J P_{njt}(\beta_m) \quad (8)$$

where $P_{njt}(\beta_m) = \frac{e^{\beta_m x_{njt}}}{\sum_{k=1}^J e^{\beta_m x_{nkt}}}$

Because the segment assignment is unknown, let H_{nm} denotes the prior probability for segment m for individual n . Then, a particularly convenient form is the multinomial logit as Eq. (9).

$$H_{nm}(\theta) = \frac{e^{z'_n \theta_m}}{\sum_{m=1}^M e^{z'_n \theta_m}}, \quad m = 1, \dots, M, \quad \theta_M = 0 \quad (9)$$

where z'_n denotes a set of observable characteristics which enter the model for class membership. The sample log-likelihood is then obtained by summing each individual's log unconditional likelihood as Eq. (10).

$$\ln L(\beta, \theta) = \sum_{n=1}^N \ln \sum_{m=1}^M H_{nm}(\theta) S_n(\beta_m) \quad (10)$$

This paper used the `lcmlogit2` package of STATA (Yoo, 2020) for implementing the latent-class logit model, which estimates the θ and β using the Expectation-Maximization (EM) algorithm. The WTP of each attribute k for each class is calculated as follows Eq. (11).

$$E(WTP_{mk}) = -E \left[\frac{\frac{\partial U_m}{\partial x_{mk}}}{\frac{\partial U_m}{\partial p_m}} \right] = -E \left[\frac{\beta_{mk}}{\beta_{m(price)}} \right] \quad (11)$$

3.2. Data

In this study, we used survey data based on conjoint analysis. Conjoint analysis refers to a method that estimates respondents' preferences based on evaluating a set of alternatives with different predefined attributes (Green and Srinivasan, 1978). Conjoint analysis has been widely used for demand analysis in the marketing area, and it was also for estimating the inconvenience of advertising loads (Shon et al., 2021) and the preference for privacy disclosure (Hann et al., 2007). Based on previous conjoint analysis research, this study estimates users' behavioral data collection preferences under the 'take-it-or-leave-it' condition based on attributes related to online platform services' hidden costs or quality factors. We choose YouTube to compare the behavioral data collection condition to attention cost because it is a representative online platform service, providing both ads-based free service and ad-free subscription service. Also, YouTube provides personalized services and advertisements with its own algorithm. It means that most YouTube users have a chance to experience the attention cost and psychological cost related to behavioral data collection. YouTube, therefore, is an appropriate service for comprehensively measuring the psychological costs to users according to differences in advertising load, ad-free subscription fees, and the terms of personal data collection.

Table 1 shows a virtual condition of using YouTube composed of four attributes used to estimate the psychological costs caused by users' behavioral data collection and advertising loads. Ads time and Monthly fee attributes are related to attention cost, and the level of attributes is based on Shon et al. (2021). In this case, the monthly fee is different from the subscription fee of YouTube Premium because this study only considers the ads-free. Shon et al. (2021) set the monthly fee for the ad-blocking service of online video from 2,000 to 8,000 KRW based on the ad-blocking services in Google Play Store and Apple App Store. The subscription fee of YouTube Premium Lite, which only provides ads-free, is 60% of YouTube premium. YouTube Premium Lite is just piloting in some countries; however, we can use it to assume the ads-free subscription fee of YouTube in South Korea more realistic. The monthly subscription fee of YouTube Premium is 10,450 KRW in South Korea, and it means that the ads-free subscription fee of YouTube in South Korea is around 6,200 KRW. So, it can be said that the attribute level of the monthly fee from 2,000 to 8,000 KRW (around \$2 ~ \$8) is reasonable. Other

attributes are related to behavioral data collection conditions. This study divided the attributes related to behavioral data collection into two major categories according to the purpose of collection referring to YouTube terms and conditions; collecting behavioral data for service improvement and/or new service development and collecting behavioral data for marketing purposes, including targeting advertisement. The level of behavioral data collection condition is two; One is ensuring users choose right when they sign-up (optional), and the other is the 'take-it-or-leave-it' choice, which means that they have to agree with the condition sign-up (necessary). We did not consider the purpose of behavioral data collection for personalized service because users can use YouTube without sign-up or log-in to avoid the personal data collected for personalized service.

Table 1. Attributes of the virtual condition of using YouTube

Attribute	Explanation and level
Ads time (advertising load)	Average non-skippable video advertising times - 5s, 10s, 15s, 20s
Monthly fee	Ads-free monthly payment fee - 2,000, 4,000, 6,000, 8,000 KRW
Behavioral data collection 1 (BC1)	Behavioral data collection for service improvement and/or new service development - condition 1 (ensuring choice right), condition 2 (take-it-or-leave-it)
Behavioral data collection 2 (BC2)	Behavioral data collection for marketing purposes, including targeting advertisement - condition 1 (ensuring choice right), condition 2 (take-it-or-leave-it)

The full factorial design matrix of the conjoint card is 32 based on four attributes for each level, and these can be combined into 496 pairs when each choice set has two alternatives. The number of whole pairs is too many to use all possible choice sets to ask. Therefore, the number of choice sets should be optimized. To choose pairs effectively, we used the D-efficiency method with STATA dcreate package (Hole, 2017). Using this method, we generated 24 choice sets with 48 alternatives, and these 24 choice sets were divided into four blocks so that each respondent could be randomly assigned one combination, including six choice sets. A sample choice set is shown in Table 2. To avoid the behavioral biases caused by the difference of understanding the behavioral data collection on YouTube, we explain the concept of behavioral data collection on YouTube before the conjoint survey.

Table 2. A sample choice set

Attributes	Condition A	Condition B
Ads time	10s	Ads free
Monthly fee	Free	4,000 KRW
Behavioral data collection 1	Take-it-or-leave-it choice	Ensuring choice right
Behavioral data collection 2	Ensuring choice right	Ensuring choice right
Choose one	()	()

Table 3. Questionnaires for factors

Factor	Questionnaire
UC	UC1. What do you think about the amount of video ads on YouTube?
	UC2. When using YouTube, how much do you think video ads cause inconvenience?
PR	PR1. It would be risky to provide personal data to online platforms.
	PR2. There would be high potential for privacy loss associated with providing personal data to online platforms.
	PR3. Personal data could be inappropriately used by using online platforms.
	PR4. Providing personal data to online platforms could involve many unexpected problems.
PB	PB1. Providing personal data to online platforms will entail benefit.
	PB2. Revealing my personal data to online platforms will help me obtain services I want.
	PB3. I think that as a result of my personal data disclosure, I will benefit from a better, more customized service.

Table 4. The result of factor analysis

Item names		Factor 1	Factor 2	Factor 3
UC	UC1	0.2044	-0.0796	0.5728
	UC2	0.1765	-0.1829	0.5836
PR	PR1	0.7630	-0.0971	0.0070
	PR2	0.8651	-0.1244	0.0799
	PR3	0.8666	-0.1468	0.0798
	PR4	0.8726	-0.1456	0.1362
PB	PB1	-0.1665	0.7222	-0.1021
	PB2	-0.1189	0.8709	-0.0558
	PB3	-0.1580	0.8331	-0.0462
Proportion of variance		0.5677	0.3962	0.1362

This study also considers individual characteristics as covariates to understand users' heterogeneity preferences of attention cost and information disclosure related to behavioral data collection. In the mixed logit model, the interaction term between the attributes and covariates suspected as the cause of heterogeneity can be used to reveal the factors that explain the heterogeneity (Hensher and Greene, 2003). Also, these covariates are used as membership variables (z'_n) in the latent-class logit model. Seven items were adopted from previous studies to measure the perceived risk and benefit of providing their personal data to online platforms (Kehr et al., 2015). Also, we designed two items to measure uncomfortable feelings because of online video advertisements. All nine questionnaires are presented in Table 3. The explanatory factor analysis is conducted to validate the measurement items, and the principal factor method and varimax rotation method are used. The result of factor analysis is summarized in Table 5. The factor analysis showed that three factors UC (uncomfortable because of ads), PR (perceived risk of personal data collection), and PB (perceived benefit of personal data collection) are retained, corresponding to three measurements. Finally, the factor scores calculated by the regression method are used in this study. YouTube usage is also considered because heavy users' perception of hidden cost and/or quality factors can be different from light users. To measure this effect, we divided respondents into two groups; The respondents who answered that they used YouTube almost every day and more than 90 minutes per day are classified as a heavy user group, and others are classified as a light user group. Lastly, the generation is also an important factor in understanding users'

online behavior. Millennials (27~42 age in 2021) and Generation Z (17~26 age in 2021) are digital natives, and especially, Generation Z never known a world without the internet (Munsch, 2021). So, the attitudes related to hidden cost and/or quality factors of free online platforms also can be different from generation. This study classified respondents into four age groups; 20s, 30s, 40s, and 50s. Individual characteristics are summarized in Table 5.

Table 5. Description of individual characteristics

Individual characteristics	Measurement
Age	20s: 2, 30s: 3, 40s: 4, 50s: 5
Usage	Heavy user: 1, light user: 0
UC	Factor score of UC (factor 3 in table 4) derived from factor analysis
PR	Factor score of PR (factor 1 in table 4) derived from factor analysis
PB	Factor score of PB (factor 2 in table 4) derived from factor analysis

The survey was conducted online by a specialized survey company among 533 YouTube users in South Korea from November 24 to November 30, 2021. Respondents who watch online videos on YouTube less than three days per week and 15 minutes per day were excluded because users who rarely use YouTube may not experience inconvenience or concern related to advertisements and personal data collection. The demographic characteristics of respondents are summarized in Table 6.

Table 6. Demographic characteristics of respondents

		# of respondents	Ratio (%)
Total		533	100.00
Sex	Male	270	50.66
	Female	263	49.34
Age (years)	20-29	132	24.77
	30-39	135	25.33
	40-49	130	24.39
	50-59	136	25.52
How many days watching YouTube per week	Almost everyday	411	77.11
	5-6 days	47	8.82
	3-4 days	75	14.07
Average time of watching YouTube per day	More than 90min	228	42.78
	75-90min	56	10.51
	60-75min	69	12.95
	45-60min	95	17.82
	30-45min	49	9.19
	15-30min	36	6.75

4. Results

4.1. Mixed logit model

Table 7 represents the result of the mixed logit model. We set all four attributes to random parameters. The mixed logit model without interaction terms showed that all means and variance of each attribute coefficient were significant at 1% level and showed a negative sign. It means that users' preference for YouTube is lower when the higher monthly price, the longer ads time, and the 'take-it-or-leave-it' condition, and there is preference heterogeneity for all attributes.

Table 7. The result of the mixed logit model

Attribute		Model 1	Model 2
Ads time	Mean of Coeff (SE)	-0.1250 (0.0143)***	-0.2042 (0.0386)***
	SD of Coeff (SE)	0.1112 (0.0182)***	0.0989 (0.0177)***
BC1	Mean of Coeff (SE)	-0.6967 (0.1056)***	-0.5608 (0.3121)*
	SD of Coeff (SE)	1.0013 (0.1689)***	0.8980 (0.1679)***
BC2	Mean of Coeff (SE)	-1.2016 (0.1187)***	-1.2814 (0.3123)***
	SD of Coeff (SE)	0.9454 (0.1734)***	0.8625 (0.1685)***
Monthly fee	Mean of Coeff (SE)	-0.0011 (0.0001)***	-0.0006 (0.0002)***
	SD of Coeff (SE)	0.0006 (0.0001)***	0.0006 (0.0001)***
Ads time x age	Mean of Coeff (SE)	-	0.0264 (0.0095)***
Ads time x usage	Mean of Coeff (SE)	-	-0.0236 (0.0212)
Ads time x UC	Mean of Coeff (SE)	-	-0.0118 (0.0129)
BC1 x age	Mean of Coeff (SE)	-	-0.0829 (0.0817)
BC1 x usage	Mean of Coeff (SE)	-	0.3078 (0.1832)*
BC1 x PR	Mean of Coeff (SE)	-	-0.3546 (0.0944)***
BC1 x PB	Mean of Coeff (SE)	-	0.2159 (0.0967)**
BC2 x age	Mean of Coeff (SE)	-	-0.0289 (0.0783)
BC2 x usage	Mean of Coeff (SE)	-	0.3922 (0.1770)**
BC2 x PR	Mean of Coeff (SE)	-	-0.1890 (0.0886)**
BC2 x PB	Mean of Coeff (SE)	-	0.3085 (0.0947)***
Monthly fee x age	Mean of Coeff (SE)	-	-0.0001 (0.0000)***
Monthly fee x usage	Mean of Coeff (SE)	-	-0.0000 (0.0001)
Observations		6,396	6,396
Log-likelihood		-1330.8438	-1289.7885
Probability of chi2		0.0000	0.0000

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

To investigate the relative importance of the condition of behavioral data collection compared to ads time (attention cost), WTP is estimated from the coefficient of the mixed logit model. The estimated WTP is summarized in Table 8. The WTP of ads time is around 216 KRW per second in a month, which means that respondents are willing to pay 216 KRW for decreasing ads 1 seconds in average. The WTP of behavioral data collection condition is around 641 KRW for service development and 1,527 KRW for marketing purposes. This result implies that the psychological cost of behavioral data collection exists, and it can vary for the purpose of personal data collection. Also, the 95 percent confidence interval showed that the psychological cost of overcoming the 'take-it-or-leave-it' condition varies among respondents. However, on average, it seems that the psychological cost of behavioral data collection based on 'take-it-or-leave-it' is not large enough to affect users' welfare in real situations compared to the psychological cost of ads time.

Table 8. WTP of the mixed logit model (model 1)

(unit: KRW)	Mean of WTP	95 percent confidence interval
Ads time	216.0774	-299.0462 – 731.201
BC1	641.4089	84.4420 – 1198.376
BC2	1527.023	-464.4601 – 3518.506

The mixed logit model with interaction terms showed that age, usage, perceived risk, and benefit are related to preference heterogeneity of each attribute. First, the mean coefficient of monthly price is negatively related to age. It can be interpreted that younger users, so-called digital natives are less reluctant to pay for digital goods than older users. Second, the mean coefficient of ads time is only positively related to age. It means that older users may be less sensitive to the advertisements loading. The interesting point is that ads uncomfortable factor is not related to preference heterogeneity of ads time. This result implies that uncomfortable with ads is not directly related to the preference for ads time, and there may be more important factors that cannot be considered in this study. Third, the mean coefficient of the condition of behavioral data collection is related to usage, perceived risk, and benefit regardless of the purpose of data collection. The interaction terms related to perceived risk and benefit showed that those factors are related to preference heterogeneity of personal data collection, which is consistent with privacy calculus theory. The interaction terms with usage imply that the heavy user group is less sensitive to the personal data collection, and it may not be common, considering the more usage means the more disclosure of behavioral data. One possible explanation for this is that heavy users may have a stronger brand relationship than light users. According to Hayes et al. (2021), the consumer-brand relationship strength can positively affect the perceived value of information disclosure. The result of the model with interaction terms implies that the individual characteristics are related to respondents' heterogeneity of psychological costs.

4.2. Latent-class logit model

To determine the number of latent classes of the model, we checked AIC (Akaike Information Criterion), CAIC (Consistent AIC), and BIC (Bayesian Information Criterion). The AIC, CAIC, and BIC as summarized in Table 9. The model did not converge when the number of latent classes was more than 6. This study chose the four classes model because its CAIC and BIC is the lowest value. The latent-class logit model results are presented in Table 10, and the estimated WTP of each attribute based on the latent-class logit model is presented in Table 11. The amount of coefficient and statistically significant attributes are varied to class. In Class 1, the coefficient of a personal data attribute is much

higher than in other groups, and the WTP of a personal data attribute is around 2,400 ~ 3,500 KRW per month, which is more than two times of average WTP in the mixed logit model. In Class 2, the coefficient of monthly price is much higher than in other groups. Therefore, the WTP of each attribute is quite lower than the mixed logit model. Also, behavioral data collection for developing a new service does not affect Class 2 users' preferences. In Class 3 and 4, the coefficient of ads time is larger than in other groups and the coefficient of behavioral data collection is insignificant. The difference between Class 3 and 4 is that Class 3's coefficient of ads time is much higher than Class 4 and Class 3's coefficient of monthly payment is lower than Class 4. Even though the WTP of ads time in Class 4 is around 290 KRW which is similar to the average of the mixed logit model, the WTP of ads time in Class 3 is more than four times that of Class 4. So, we can interpret that the respondents in Class 1 are very sensitive to behavioral data collection compared to advertising loads, respondents in Class 2 have a low preference for paid services, and respondents in Class 3 are very sensitive to advertising load (attention cost) but not sensitive to provide behavioral data. Respondents in Class 4 seem that they also are sensitive to ads time, but they do not want to pay much as Class 3 for ads-free service. The probability that respondents are classified as Class 1, 2, 3, and 4 is 32.9%, 47.0%, 6.6%, and 13.4% respectively.

Table 9. Result of AIC, CAIC, and BIC

# of classes	Log-likelihood	# of parameters	AIC	CAIC	BIC
2	-1338.107	14	2704.214	2778.114	2764.114
3	-1311.739	24	2671.477	2798.162	2774.162
4	-1252.116	34	2572.233	2751.702	2717.702
5	-1234.792	44	2557.585	2789.84	2745.84

Table 10. Result of latent-class logit model

Attribute	Class 1	Class 2	Class 3	Class 4
Ads time	-0.0448 (0.0131)***	-0.1239 (0.0218)***	-0.5312 (0.2251)	-0.2887 (0.0560)***
BC1	-0.8931 (0.1243)***	-0.3951 (0.1812)***	-0.2268 (0.4259)	0.1576 (0.2804)
BC2	-1.3022 (0.1329)***	-1.1155 (0.2127)**	-0.1087 (0.4070)**	0.3605 (0.3430)
Monthly fee	-0.0004 (0.0000)***	-0.0020 (0.0003)***	-0.0004 (0.0001)***	-0.0010 (0.0002)***
Age	0.0703 (0.2348)	0.4926 (0.2081)**	-0.3344 (0.2770)	-
Usage	-0.0762 (0.4252)	0.2014 (0.3849)	0.6429 (0.5247)	-
UC	-0.4656 (0.2878)	-0.2046 (0.2628)	0.2548 (0.3578)	-
PR	0.8011 (0.2819)***	0.4765 (0.2689)*	0.4749 (0.3168)	-
PB	-0.9611 (0.3139)***	-0.8492 (0.2960)***	-0.1978 (0.3602)	-
Constant	0.9439 (0.9567)	-0.2665 (0.8256)	0.1064 (1.0321)	-
Probability of Class	32.94%	47.02%	6.64%	13.40%
Log-likelihood	-1252.547			
Observations	6,396			

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

Table 11. WTP of latent-class logit model

	(unit: KRW)	Mean of WTP	95 percent confidence interval
Class 1	Ads time	121.1193	64.8873 – 177.3512
	BC1	2416.674	1754.591 – 3078.757
	BC2	3523.78	2672.997 – 4374.564
Class 2	Ads time	63.4513	43.8182 – 83.0844
	BC1	202.2914	-14.0761 – 418.6590
	BC2	571.1924	379.4126 – 762.9722
Class 3	Ads time	1375.20	881.1945 – 1869.203
	BC1	587.0742	-1325.861 – 2500.009
	BC2	281.4173	-1900.602 – 2463.437
Class 4	Ads time	293.4402	232.0284 – 354.852
	BC1	-160.1370	-682.0099 – 361.7359
	BC2	-360.4216	-957.5666 – 224.7234

The membership coefficient showed that respondents with higher perceived risk and lower perceived benefit of behavioral data collection have a higher probability of belonging to Class 1 than Class 4 (reference class). Also, older respondents with higher perceived risk and lower perceived benefit of behavioral data collection have a higher probability of belonging to Class 2 than Class 4. Furthermore, there was no significant difference between Class 3 and Class 4 regarding membership coefficient. One of interesting points is that Class 2 is similar to Class 1 in terms of membership coefficient of perceived risks and benefits. However, their WTP is much lower than Class 1. It can be interpreted that respondents who are not used to paying for online services have a low WTP for overcoming the 'take-it-or-leave-it' condition even though they do not prefer that condition. In other words, they seem to accept the 'take-it-or-leave-it' condition as an exchange for free service. On the other hand, in the case of Class 3 and Class 4, the WTP of personal data collection was estimated to be low because they do not perceive the personal behavioral data collection as particularly harmful even though they are willing to pay for better usage conditions. It means that there are three distinct segments in terms of accepting the 'take-it-or-leave-it' condition of personal behavioral data collection; accepting the condition due to the lack of a sufficient option (Class 1), accepting the condition in exchange for free service (Class 2), accepting the condition because it does not matter (Class 3 and Class 4). These different attitudes toward the 'take-it-or-leave-it' condition imply that it may be inappropriate to enforce the identical condition for all users. In this regard, it can be said that the prohibition of the dominant online platform from collecting personal data excessively without sufficient choice is a meaningful remedy to protect consumers' welfare.

5. Conclusion

5.1. Summarize and implications

The mixed logit model and latent-class logit model showed preference heterogeneity related to behavioral data collection. Moreover, the latent-class logit model showed that the psychological costs of consenting to the collection of behavioral data collection for each segment could be very different. It implies that it may be inappropriate to enforce identical conditions for all users. Considering the preference heterogeneity and the high WTP of privacy-sensitive respondents, we suggest the need to develop options with appropriate trade-offs so that privacy-sensitive users with high WTP can choose to reject the behavioral data collection. Similarly, in the case of advertisements where users also have different preferences, the ads-free option with a monetary payment, such as YouTube Premium, has already been introduced. In this approach, online platforms could collect users' behavioral data who have low WTP in exchange for providing free service while providing sufficient options for users who have high WTP. The result that the perceived risk and benefit of providing personal data affect the psychological cost of behavioral data collection implies that the privacy calculus theory is still a meaningful framework for explaining the heterogeneity of users' attitudes. It implies that users could manage their privacy choice based on expected loss and gain of information disclosure when the options with appropriate trade-offs are ensured. However, their decision also can be biased by behavioral biases. Therefore, sufficient information related to the purpose and scope of collecting and processing personal behavioral data should be provided to avoid behavioral biases when choosing the option. Also, it seems necessary to consider ways to minimize the users' perceived risks of providing personal data, such as enhancing data security for online platforms that need to provide sufficient options when collecting personal data from users. Lastly, our result showed that the psychological cost of consenting to the behavioral data collection varies from the purpose of collecting data. It can be interpreted that the perception that the scope of behavioral data may vary to the purpose of data collection affects the users' perceived risk and benefit. Considering this, regulators or lawmakers should consider the purpose and scope of behavioral data collection when introducing or implementing the regulation. Too stringent regulation on data collection may negatively impact social welfare by reducing the efficiency of data-driven business models.

5.2. Limitations and future research

Our study is different from previous research by quantifying the different psychological costs of personal data collection with the attention cost, however, there are several limitations. First of all, our study is just focused on users in South Korea. Previous studies noted that the attitude related to privacy and personal data collection could vary according to nationality because culture, the general perception, and the degree of privacy regulation can differ (Winegar and Sustein, 2019; Prince and Wallsten, 2021; Alfnes and Wasenden, 2022). So, future research could provide meaningful implications by conducting cross-country analysis considering the difference of privacy regulatory environment. Second, except for the generation (age) variable, meaningful variables related to the attention cost and preference of free service were not included as covariates in this analysis. So, further research could explain the individual's heterogeneity of the psychological costs in more detail by including the variables related to attention cost and preference of free service.

References

- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221), 509-514.
- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2020). Secrets and likes: the drive for privacy and the difficulty of achieving it in the digital age. *Journal of Consumer Psychology*, 30(4), 736-758.
- Acquisti, A., John, L. K., & Loewenstein, G. (2013). What is privacy worth?. *The Journal of Legal Studies*, 42(2), 249-274.
- Alfnes, F., & Wasenden, O. C. (2022). Your privacy for a discount? Exploring the willingness to share personal data for personalized offers. *Telecommunications Policy*, 46(7), 102308.
- Baek, T. H., & Morimoto, M. (2012). Stay away from me. *Journal of advertising*, 41(1), 59-76.
- Barth, S., & De Jong, M. D. (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.
- Bundeskartellamt. (2019). Facebook, Facebook, Exploitative business terms pursuant to Section 19(1) GWB for inadequate data processing. Case Summary, B6-22/16.
- Choi, B., Wu, Y., Yu, J., & Land, L. P. W. (2018). Love at first sight: The interplay between privacy dispositions and privacy calculus in online social connectivity management. *Journal of the Association for Information Systems*, 19(3), 4.
- CMA (2020). Online platforms and digital advertising market study.
- Cooper, J. C. (2012). Privacy and antitrust: Underpants gnomes, the first amendment, and subjectivity. *Geo. Mason L. Rev.*, 20, 1129.
- Costa-Cabral, F., & Lynskey, O. (2017). Family ties: the intersection between data protection and competition in EU Law. *Common Market L. Rev.*, 54, 11.
- Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information systems research*, 17(1), 61-80.
- Evans, D. S. (2013). Attention rivalry among online platforms. *Journal of Competition Law & Economics*, 9(2), 313-357.
- Ezrachi, A., & Robertson, V. H. (2019). Competition, market power and third-party tracking. *World Competition*, 42(1).
- Gal, M. S., & Rubinfeld, D. L. (2016). The hidden costs of free goods: Implications for antitrust enforcement. *Antitrust Law Journal*, 80(3), 521-562.
- Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: issues and outlook. *Journal of consumer research*, 5(2), 103-123.
- Hann, I. H., Hui, K. L., Lee, S. Y. T., & Png, I. P. (2007). Overcoming online information privacy concerns: An information-processing theory approach. *Journal of management information systems*, 24(2), 13-42.
- Hayes, J. L., Brinson, N. H., Bott, G. J., & Moeller, C. M. (2021). The influence of consumer–brand relationship on the personalized advertising privacy calculus in social media. *Journal of Interactive*

Marketing, 55, 16-30.

Hensher, D. A., & Greene, W. H. (2003). The mixed logit model: the state of practice. *Transportation*, 30(2), 133-176.

Hermes, S., Clemons, E. K., Wittenzellner, D., Hein, A., Böhm, M., & Krcmar, H. (2020). Consumer Attitudes towards Firms that Monetize Personal Information: A Cluster Analysis and Regulatory Implications. In *PACIS* (p. 29).

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata journal*, 7(3), 388-401.

Hole, A. R. (2017). DCREATE: Stata module to create efficient designs for discrete choice experiments.

Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, 25(6), 607-635.

Kim, D., Park, K., Park, Y., & Ahn, J. H. (2019). Willingness to provide personal information: Perspective of privacy calculus in IoT services. *Computers in Human Behavior*, 92, 273-281.

Kimmel, L., & Kestenbaum, J. (2014). What's up with WhatsApp: a transatlantic view on privacy and merger enforcement in digital markets. *Antitrust*, 29, 48.

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security*, 64, 122-134.

Lerner, A. V. (2014). The role of 'big data' in online platform competition. Available at SSRN 2482780.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

Munsch, A. (2021). Millennial and generation Z digital marketing communication and advertising effectiveness: A qualitative exploration. *Journal of Global Scholars of Marketing Science*, 31(1), 10-29.

Newman, J. M. (2016). Antitrust in zero-price markets: applications. *Wash. UL Rev.*, 94, 49.

OECD (2013). *Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value*. OECD Digital Economy Papers, No. 220, OECD Publishing, Paris.

Prince, J., & Wallsten, S. (2021). Empirical Evidence of the Value of Privacy. *Journal of European Competition Law & Practice*, 12(8), 648-654.

Reyna, A. (2018). The psychology of privacy-what can behavioral economic contribute to competition in digital markets?. *International Data Privacy Law*, 8(3), 240-252.

Robertson, V. H. (2020). Excessive data collection: privacy considerations and abuse of dominance in the era of big data. *Common Market Law Review*, 57(1).

Savage, S. J., & Waldman, D. M. (2015). Privacy trade-offs in smartphone applications. *Economics Letters*, 137, 171-175.

Shon, M., Shin, J., Hwang, J., & Lee, D. (2021). Free contents vs. inconvenience costs: Two faces of online video advertising. *Telematics and Informatics*, 56, 101476.

Stucke, M. E., & Ezrachi, A. (2016). When competition fails to optimize quality: A look at search engines. *Yale JL & Tech.*, 18, 70.

Stuke, M. E. (2018). Should we be concerned about data-opolies? 2 *GEO. L. TECH. REV.* 275.

- Train, K. E. (2009). Discrete choice methods with simulation. Cambridge university press.
- Which? (2021). Value of the Choice Requirement Remedy. Research Report September 2021.
- White, T. B., Zahay, D. L., Thorbjørnsen, H., & Shavitt, S. (2008). Getting too personal: Reactance to highly personalized email solicitations. *Marketing Letters*, 19(1), 39-50.
- Winegar, A. G., & Sunstein, C. R. (2019). How much is data privacy worth? a preliminary investigation. *Journal of Consumer Policy*, 42(3), 425-
- Wu, T. (2018). Blind spot: The attention economy and the law. *Antitrust LJ*, 82, 771.
- Yoo, H. I. (2020). *lclogit2*: An enhanced command to fit latent class conditional logit models. *The Stata Journal*, 20(2), 405-425.
- Zuiderveen Borgesius, F. (2014). Behavioural Sciences and the Regulation of Privacy on the Internet. Draft chapter for the book 'Nudging and the Law-What can EU Law learn from Behavioural Sciences.