Consumer Attitudes towards Firms that Monetize Personal Information: A Cluster Analysis and Regulatory Implications

Completed Research Paper

Sebastian Hermes
Technical University of Munich
Munich, Germany
sebastian.hermes@tum.de

Eric K. Clemons
The Wharton School
Philadelphia, United States
clemons@upenn.edu

David Wittenzellner
Technical University of Munich
Munich, Germany
david.wittenzellner@tum.de

Andreas Hein
Technical University of Munich
Munich, Germany
andreas.hein@tum.de

Markus Böhm
Technical University of Munich
Munich, Germany
markus.boehm@tum.de

Helmut Krcmar
Technical University of Munich
Munich, Germany
helmut.krcmar@tum.de

Abstract

The EU is seeking to develop new regulatory frameworks for online privacy. This entails a complex set of tradeoffs, since regulatory policy must be informed by consumers' preferences, and if regulatory policy contravenes consumer preferences, regulators might need to explain the hidden sources of harm to consumers. To increase our understanding of consumers' attitudes towards firms that monetize privacy, we surveyed 1693 individuals from Denmark, France, Germany, the UK, and the USA about Google. Our cluster analysis confirms the privacy paradox – although consumers disapprove of Google's practices, they still use it – in four out of five clusters but indicates two different explanations: some consumers cannot locate a viable alternative, whereas others lack the information needed for their privacy calculus. We explore regulatory implications and draw upon the theory of newly vulnerable markets to discuss whether market entry may be feasible.

Keywords: Consumer attitudes towards privacy, informed consent, privacy paradox, privacy regulation, theory of newly vulnerable markets

Introduction

We are reporting on a set of consumer surveys that seek to examine consumers' attitudes towards firms that monetize consumers' personal information by leveraging data mining and personalized ads. The surveys assessed consumers' knowledge of Google's current business practices, as well as their approval/disapproval of those practices, independent of their belief that those practices currently incurred. This allowed us to determine *implicit informed consent;* individuals who both *knew of* and *approved of* these practices. We also investigated their willingness to switch to alternative *safer* services, that is, services that do not violate privacy or monetize private information. We studied their

degree of regulatory satisfaction, including the belief that regulators were taking adequate steps to inform consumers about potential privacy risks online and to protect them from these risks. We saw very little differences between consumers in France, Germany, the UK, and the USA. When compared with previous studies (Clemons et al. 2014; Clemons and Wilson 2015), we found a slight increase in *informed consent*; this was not due to an increase in the number of consumers who approved of Google's practice, rather by a slight increase in the number of consumers who are now aware of Google's practices

Others researchers have established the presence of real, measurable harm from Google, suggesting that the current levels of regulatory protection may be inadequate (e.g. Edelman 2015; Edelman and Geradin 2016; European Commission 2017; European Commission 2018; European Commission 2019; European Parliament 2019; Iacobucci and Ducci 2019; Zhu and Liu 2018). Our study contributes to the design of regulatory policies to improve consumer welfare and decrease the present and future harm created by firms that monetize personal information that fall outside existing regulatory frameworks.

Thereby, we were guided by the following research question: *What clusters of informed consent exist and how should regulators intervene?* We surveyed 1693 individuals in Denmark, France, Germany, the UK, and the USA regarding Google. Our cluster analysis confirms the privacy paradox – although consumers disapprove of Google's practices, they still use it – in four out of five clusters but indicates two different explanations: some consumers cannot locate a viable alternative, whereas others lack the information needed for their privacy calculus.

Moreover, we draw on the theory of newly vulnerable markets (NVM) to discuss whether market entry may be feasible. Our results indicate that market entry is feasible for cluster 5 (willing to use and somewhat willing to pay for safe alternative and disapproving of Google's practices), rather difficult for cluster 2 (approving Google's practices and willing to grant it a GDPR waiver but somewhat willing to pay for safe alternative), rather difficult for cluster 1 and 3 (willing to use but unsure about paying for safe alternative and disapproving Google's practices), and completely unfeasible for cluster 4 (unwilling to use and pay for a safe alternative although disapproving Google's practices). We discuss which form of regulatory intervention might be necessary for each cluster. We argue that no intervention might be necessary for cluster 5 and 2, whereas clusters 1 and 3 and cluster 4 need different interventions in form of being better informed through regulation. Cluster 1 and 3 need to be better informed about the logic and advantages of safe alternatives to cope with their disapproval of Google's practices and their undecidedness to pay for a safe alternative. Cluster 4 needs to be better informed in two areas. First, about the hidden costs of apparently free services as it indicates not to be well informed about Google's practices. Second, about the logic and advantages of safe alternatives to cope with its disapproval of Google's services and its unwillingness to use and pay for safe alternatives. In contrast, cluster 5 is willing to use and somewhat willing to pay for safe alternatives and, therefore, might render regulatory intervention unnecessary. Since cluster 2 provides, on an aggregated level, informed consent for Google's practices and is willing to use and pay for safe alternative, no regulatory intervention might be necessary for this cluster neither.

This paper makes two contributions. The first is *practical*: The EU is seeking to develop new regulatory policies with multiple objectives. These include providing adequate (privacy) protections for consumers and providing a fair and level playing field that protects competition and limits the abuses created by firms that monetize personal information. Developing new regulatory policies entails a complex set of tradeoffs since regulatory policy must be informed by consumers' preferences, and if regulatory policy contravenes consumer preferences, it may be necessary for regulators to explain the hidden sources of harm to consumers. This study provides an input to the EU regulatory process. The second is *theoretical*: The privacy paradox indicates that consumers' online behavior often contradicts their expressed preferences for safe, secure, and private online interactions. We observed that while consumers express a preference for safer alternatives, they may be (a) unaware of the hidden costs of using existing Google offerings, (b) unable to find potentially safer alternatives with all the functionality they now obtain from Google, or (c) unwilling to pay for safer alternatives. This last point suggests that while consumers may have a real preference for privacy, the value placed on privacy may be quite low.

Theoretical Underpinnings

Monetization of Personal Information: Price Discrimination, Data Mining, and Personalized Advertising

Service providers who are well informed about their customers can engage in price discrimination (Phlips 1983). That means service providers are using customer information (Weking et al. 2018) to determine which type of customer is willing to pay which type of price for the service offered. As illustrated by Clemons and Wilson (2015), airlines, for example, seek to determine whether a customer is a business or leisure traveler to charge higher prices to the former since they assume that business travelers have to travel and do not need to cover the costs themselves. Thus, under an situation of complete price discrimination, service providers could exactly charge the customer's surplus.

Since service providers are aiming to maximize profits, they are looking to leverage new information technologies such as data mining to move towards complete price discrimination (Clemons et al. 2014). We refer to data mining as the systematic application of statistical methods to large data sets to identify new links and trends. Data mining, therefore, helps the service provider to move beyond identifying sheer customer types towards identifying specific needs and characteristics of each customer at a particular point in time. Although firms do not have sufficient data for complete price information, the work of Shiller (2014) indicates how big data in combination with data mining is approximating complete price discrimination.

Online advertising companies, such as Google, have specialized in creating detailed profiles of consumers and can help service providers to charge individual prices (Clemons and Wilson 2015; Zuboff 2019). Due to Google's vast amount of services such as search, video streaming, email, maps, or cloud services, Google can mine various forms of online and offline behavior and integrate different data sources into one comprehensive consumer profile (Clemons 2018). In addition, Google even engages in illegal conduct and violates consumers' privacy to improve further its personal profiles (FTC 2019; Porter 2018). Due to Google's comprehensive consumer profiles, service providers (advertisers) draw on Google ads because it allows them "to know not only *who* wants *what*, but *how badly they want it*, and by inference, *how much they are willing to pay*" (Clemons and Wilson 2015). That is, while consumers may be anonymous to the advertiser before they click an ad, they can be associate by the advertiser with specific attributes and characteristics after they clicked an ad. Hence, consumers are not anonymous anymore and advertiser can determine individual prices. Consequently, consumers might associate themselves to certain behaviors, attributes, or characteristics with which they would not want to be associated or would not want advertisers to know, creating a source of consumer harm.

The Privacy Paradox

The privacy paradox refers to consumers indicating that they have high levels of privacy concerns, while simultaneously engaging in unsafe practices and ignoring the risk of losing control over their personal data (Gerber et al. 2018). In other words, a contradiction between consumers' stated preferences and their actual behavior (Kokolakis 2017). Many consumers express concerns about their data, their desire to protect it, and their desire to maintain control over its access. These same consumers, however, disclose a variety of personal data, often without reviewing the privacy policy of the service provider (Chakraborty et al. 2013). There are three possible explanations:

Privacy calculus theory postulates that individuals perform a calculus between the expected cost of loss of privacy and the potential gain of their unsafe behavior. Their final behavior is determined by the outcome of this privacy calculus (Dinev et al. 2006; Dinev and Hart 2006; Xu et al. 2011). Thus, it is assumed that individuals consciously and rationally decide to disclose personal information when potential gains exceed expected losses.

However, research in behavioral economics demonstrates that human decision-making is affected by **cognitive biases and heuristics** (Acquisti et al. 2012; Adjerid et al. 2018). These authors claim that privacy decisions are often affected by the same biases and heuristics.

Moreover, **bounded rationality and incomplete information** are also significant. Human decision-making, even executive decision-making, is seldom fully informed and seldom fully rational (Simon 1997). Individuals make privacy decisions with incomplete information about risks and benefits and seldom perform complex calculations, enabling them to make rational decisions (Acquisti and Grossklags 2005).

Informed Consent

The concept of informed consent can mostly clearly be seen in the domain of medical treatment. Its underlying idea is that patients have to be informed and approve of the suggested medical treatment (Schenker et al. 2011). Hence, informed consent provides patients some level of control over medical decision-making by requiring that they be fully informed about the risks and benefits of treatment, delaying treatment, and alternative treatment. Obtaining informed consent is necessary for doctors to avoid litigation if medical complications arise during treatment. Informed consent is thus a legal construct, with clearly defined terms and conditions. Terms and conditions are essential to informing users in the process of obtaining informed consent. They provide the description of future actions that cannot be performed without users' prior agreement. Understanding the terms and conditions is critical.

Informed consent is also significant when users are offered unsafe products or products with a known degree of risk, such as scuba instruction. Before engaging in such activities, the instructor ensures that users are aware of and accept the risks. This is critical in protecting service providers in case of an accident.

Understanding the terms and conditions is likewise critical in agreements between users and online service providers. Individuals must fully understand what they are agreeing on when considering to submit their personal data to the processing purposes of others (Luger et al. 2013). Something as simple as ticking the box to indicate that a user accepts the terms and conditions must be based on a clear understanding of what exactly is going to be done. Without this understanding, the user is not *truly* informed, and the consent offered cannot reasonably be called informed consent.

However, informed consent has previously not played a significant role in privacy research in the information systems literature, despite the fact that that informed consent is a core aspect of data protection. Exceptional is the work of Clemons et al. (2014) and Clemons and Wilson (2015). The General Data Protection Regulation (GDPR) explicitly requires companies to obtain informed consent before collecting and processing personal data and to be explicit about what will be done with the data (European Parliament and the Council of the EU 2016). Often, obtaining consent is usually not much more than the ritual of scrolling down business conditions and checking the box in order to access a digital service. Most users are well known to neither completely read privacy policies nor completely understand them (Chakraborty et al. 2013; Zuboff 2019). Nonetheless, the GDPR continues to adhere to the principle of informed consent. Despite this discrepancy, research has not investigated how much users know about the ways that companies use their data at an atomic, function-by-function level and whether the approve of these practices or not.

We refine the concept of informed consent used by Clemons et al. (2014) by introducing the concept of *implicit informed consent*, which is implicitly given by users who are aware that their online service provider performs a specified activity using their data *and* approve of that use. Users who accept the terms and conditions without fully understanding the actions that an online service provider is allowed to perform cannot be considered to have provided implicit informed consent. Contrarily, users who are fully informed about the actions of an online service provider and approve of them indeed have provided implicit informed consent. We believe that a user who agrees to a policy, is aware of it, and does not approve of it but accepts only because she cannot locate an acceptable alternative has not provided implicit informed consent.

Theory of Newly Vulnerable Markets

The theory of newly vulnerable markets explains changes in competition, in which new competitors are capable of attacking stronger incumbents, even when incumbents enjoy reputational advantages,

economies of scale, and superior market share (Clemons 2018; Clemons et al. 2002). The NVM theory builds on the theory of contestable markets (Baumol 1982), which posits that in monopolies or oligopolies that are under continuous threat of new entry, prices are similar to perfect competition to eliminate the incentives of new entrants to enter the market. The three components of the NVM theory are as follows.

Newly easy to enter: This occurs when technological or regulatory changes reduce entry barriers or when changes in consumer preferences weaken the competitive advantage of previously dominant firms.

Attractive to attack: This implies the presence of a strong customer profitability gradient, that is, extreme differences in profitability across customers. This usually occurs when existing firms in an industry charge the same prices to all customers, even when customers differ substantially in terms of willingness to pay for goods and services, or in their costs to serve. New entrants generally choose to target the most profitable customers.

Difficult to defend: This occurs when the incumbents experience barriers in changing their strategy to imitate the attacker's strategy. This may be caused by asymmetric regulation that favors new entrants, existing contracts that require continued service to less attractive customer segments, or other factors that restrict incumbents' ability to adopt the practices of new entrants.

We will explore whether changes in customers' attitudes towards privacy create the conditions for newly vulnerable markets.

Methodology

Data Collection. We administered an online survey to consumers in Germany, the UK, the USA, France, and Denmark, exploring their knowledge and approval of Google's business practices. We chose Google as an example for monetizing and violating privacy (Porter 2018; Zuboff 2019). We obtained survey results from 1693 individuals, 364 in Denmark, 333 in France, 350 in Germany, 321 in the UK, and 325 in the USA[1]. The survey instrument was an extension of the survey that we previously used (Clemons et al. 2014; Clemons and Wilson 2015), with questions added to reflect some of the uses of private information that have emerged since our prior studies. To ensure that the questions were clear and that the subjects were able to respond to them, we conducted a pretest. These pretest results were not included in the final data. Participants were ensured anonymity.

Survey Measurements. Participants were asked to answer questions on their personal attitudes towards Google, safer alternatives, and regulatory protections. To eliminate sequence effects, the order of questions was shuffled between the subjects. Our first construct, implicit informed consent, was measured by asking participants to indicate their awareness and degree of approval of 24 tracking and privacy-related practices of Google.

The first subconstruct explored simple tracking, e.g., "Does Google track your GPS location history?" The second was tracking integration across services, e.g., "Does Google create a composite profile involving your search history, and the content of your emails and texts?" Finally, we addressed implications, e.g., "Does Google maintain enough information about you to infer your political affiliation?" Each awareness question was measured on a 3-point Likert scale (They do, do not know, they do not) and each consent question on a 5-point Likert scale (Strongly object to strongly approve).

For the analysis, we computed an awareness score and a consent score. A participant's awareness score was incremented by one if he/she answered an awareness question with "They do" and decremented if they answered, "They don't." Similarly, the consent score was incremented with the answer "Agree" or "Strongly Agree" and decremented with the answer "Object" or "Strongly Object." Consequently, the awareness and consent scores have a theoretical range of -24 to 24.

The second area of interest: the attitudes towards alternative online service providers covered four

[1] Our sample sizes were large enough to be significant, and to allow a significant participation in enough demographic groups to ensure that we are representative of each nation surveyed. A copy of the survey instrument and participant demographics will be provided in the final paper.

questions. First is whether participants would grant startups a GDPR waiver if they offered the same service as Google; second, whether they would be more comfortable using an EU alternative; third, whether they would use a safe alternative in preference to Google; and finally, whether they would pay for a safe alternative in preference to Google. The answers were provided on a 5-point Likert scale.

The third area of interest covered participants' history of usage of Google Search, Gmail, Maps/Waze, and alternatives. Participants indicated which applications they routinely use. Fourth, we explored participants' attitudes towards the level of regulatory protection they currently receive. Moreover, we evaluated their satisfaction with the degree to which regulation ensured that they were aware of how online data is used and the potential forms of harm, as well as the degree to which regulators protected them from online harm caused by service providers' use of data. Again, we calculated a composite score. This regulatory satisfaction score was incremented if the participant answered "Agree" or "Strongly Agree." It was decremented if the participant answered "Disagree" or "Strongly Disagree." This enabled a theoretical range of −3 to 3.

Finally, we asked participants sociodemographic questions and also asked them to indicate their level of technological sophistication and whether they have encountered online privacy violations.

Data Analysis. We first report simple statistics, aggregate levels of Awareness of Google's activities, and aggregate levels of Acceptance of Google's activities. We found very little variation across countries, although the levels of acceptance were slightly higher in the USA and slightly lower in Germany. Since there was little variation, we do not separate our results by subjects' nationality. We then report simple cross-tabulations, showing the percentages of subjects who were aware and approved of Google's practices. This is our measure of implicit informed consent.

We next applied exploratory factor analysis to merge potentially similar items into a smaller set of first-order constructs and to assess whether the awareness, consent, and regulatory items could be represented by one composite score for each construct. We used principal axis factor as extraction method and varimax as rotation method. The Gorsuch–Nelson interpretation of the scree plot suggests the extraction of three factors. Each of the three resulting factors distinctively represented awareness, consent, or regulatory satisfaction, as they only exhibited relevant loadings on those respective questions (see Appendix). Each item loaded only on one factor (cutoff = 0.3).

Our clustering method used the following variables: awareness, consent, GDPR waiver for startups, usage of a European alternative, usage of clean alternative, paying for clean alternative, and regulatory satisfaction. In order to determine the number of clusters, we ran multiple statistics, such as silhouette plots, the Krzanowski–Lai index, and the gap statistic. Most statistics suggested two, three, or five clusters. We used five clusters to increase the explanatory power and to account for reasonable content-based segmentation. The data was then clustered using a hierarchical k-means approach, which is also known as two-step clustering. In the first step, the data was clustered hierarchically. This process was stopped when five clusters had been formed. The centers of these hierarchical clusters were then used as a starting point for the k-means algorithm. The distances were calculated based on the Euclidean distance.

Results

Descriptive Analysis of Awareness and Consent

The descriptive analysis reveals the following insights about awareness of Google's practices (- 24 not aware at all to +24 completely aware). More people were found to be aware (0 to 24) than unaware (−24 to 0). The majority has a score of 0 (they are aware of 50% and unaware of 50%).

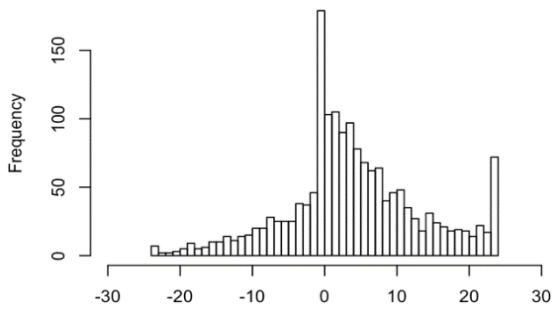

Figure 1: Histogram Awareness

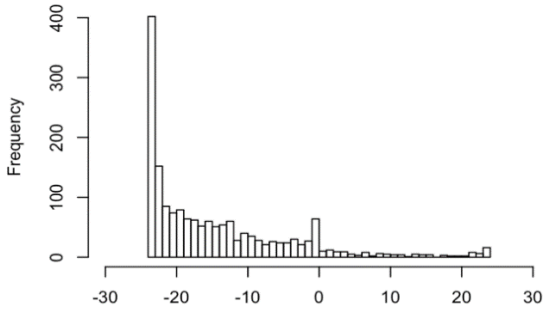

Figure 2: Histogram Consent

Regarding consumers' level of approval (- 24 completely disapprove to 24 completely approve), we obtained the following results: the far-left bar indicates that most individuals are providing no consent at all; the majority of individuals are not providing consent (-24 to 0).

In summary, people have positive levels of awareness and very low levels of approval of Google's practices.

Descriptive Analysis of Informed Consent

Table 1 presents the cross-tabulations of awareness and approval of 24 Google activities. The three highest levels of informed consent were for tracking GPS location history (10%), retaining and using search history (9,7%), and controlling search to provide better placement for Google's own offerings (6,8%). In part, these items scored higher on informed consent because more subjects were aware of them. The four lowest levels of informed consent were for Google's permitting data to leave the subject's home country and to use it in ways not level in the home country (2,5%); determining the user's religious affiliation (2,5%); estimating the user's net worth (2,5%), and retaining this data for as long as Google believes it will be valuable (2,4%).

We found that only 6 out of 24 of Google's activities had levels of informed consent of 5% or above. No activity had informed consent level above 10%. This cannot be explained solely by the levels of awareness. Users generally disapprove of almost all of Google's activities.

Table 1: Degree of Informed Consent

Category	ID[1]	Business conduct	Informed Consent
Individual Practices	Q19/Q20	Does Google track, retain, and use your GPS location history? If they did, would you care?	10,00%
	Q9/Q10	Does Google track, retain, and use your search history? If they did, would you care?	9,70%
	Q11/Q12	Does Google track, retain, and use the contents of the emails sent to a Gmail account? If they did, would you care?	3,90%
	Q17/Q18	Does Google's technology allow it to track, retain, and use the actual words in the contents of the voice messages left on an Android device? If they did, would you care?	3,90%
	Q13/Q14	Does Google track, retain, and use the contents of the emails sent from a Gmail account? If they did, would you care?	3,80%
	Q21/Q22	Does Google data mine your contact list and your calendar? If they did, would you care?	3,80%
	Q15/Q16	Does Google track, retain, and use the contents of text messages sent to or from an Android device? If they did, would you care?	3,30%
	Q23/Q24	Does Google data mine your doc files and spreadsheets? If they did, would you care?	2,90%
Integration of Practices	Q25/Q26	Does Google integrate information from numerous sources to create a composite profile of each of their hundreds of millions of users? If they did, would you care?	5,60%
	Q37/Q38	Does Google use the information it collects to know your daily routines, including where you are now and where you are going to be next? If they did, would you care?	4,90%
	Q27/Q28	Does Google use the information it collects to understand your health? If they did,	4,30%

Category	ID[1]	Business conduct	Informed Consent
		would you care?	
	Q33/Q34	Does Google use the information it collects to determine your political affiliation? If they did, would you care?	4,00%
	Q35/Q36	Does Google use the information it collects to know your daily routines, including the names of the people you see and where you see them? If they did, would you care?	3,30%
	Q29/Q30	Does Google use the information it collects to understand your sexual orientation? If they did, would you care?	2,80%
	Q31/Q32	Does Google use the information it collects to determine your religious affiliation? If they did, would you care?	2,50%
	Q39/Q40	Does Google use the information it collects to estimate your net worth? If they did, would you care?	2,50%
Implications of Practices	Q56/Q57	Does Google use its control over search to promote its own products, perhaps by placing them at the top of search results? If they did, would you care?	6,80%
	Q58/Q59	Does Google use its control over Android devices to promote its own products, perhaps by ensuring them preferred locations on the screens of smartphones and tablets? If they did, would you care?	5,30%
	Q45/Q46	Does Google share the information it collects in ways that allow a travel services provider to infer how urgently you need to travel and what you are willing to pay for a specific flight or hotel? If they did, would you care?	5,00%
	Q43/Q44	Does Google share the information it collects in ways that allow a merchant to make inferences about your willingness to pay for a specific item? If they did, would you care?	3,90%
	Q47/Q48	Does the information Google shared remain with a service provider even after you have cleared your search history or your GPS history? If they did, would you care?	3,80%
	Q41/Q42	Does Google share the information it collects in ways that allow a medical service provider to make inferences about your health? If they did, would you care?	2,80%
	Q51/Q52	Does Google permit the information it gathers to leave your home country and to be used in ways that might not be permitted in your home country? If they did, would you care?	2,50%
	Q49/Q50	Does the information Google shared remain with a service provider even after you have requested that information be removed from Google search results under your Right to be Forgotten Online? If they did, would you care?	2,40%

[1] *= ID of Informedness Question / ID of Consent Question*

Cluster Analysis

Table 3 presents the quantitative results of the cluster analysis, presenting construct values and demographical differences by cluster. **Cluster 1** accounts for the highest values regarding the usage of alternatives to Google, being aware of Google's practices and capabilities and disapproving of these practices and capabilities. Similar to the other clusters, cluster 1 is unsure about being more comfortable about using a European alternative, unsure about providing startups a GDPR waiver, willing to use but unsure to pay for a safe alternative, and indicates low frequencies of using alternatives to Google. Moreover, the cluster demonstrates strong dissatisfaction with regulators.

Cluster 2 distinguishes itself from the first one, particularly regarding its approval of Google's practices and capabilities. The second cluster is actually the only cluster that approves rather than disapproves of Google's practices. Interestingly, the cluster still indicates willingness to use and to pay for a safe alternative. However, the cluster is unsure or does not care if the alternatives stem from Europe. Moreover, the cluster shows the same frequencies of using Google services as the other clusters and is not sure whether or not to provide a GDPR waiver to startups. Similar to cluster 5, cluster 2 is largely satisfied with regulators.

Table 2: Construct characteristics by cluster

Variable	Measure	Cluster 1	2	3	4	5
Size	Count	345	205	459	314	370
Age	Selection	45,14	37,78	46,44	42,26	44,16
Sex	Selection	1,47	1,38	1,50	1,58	1,57
Google search	Frequency of selection	90%	83%	88%	87%	89%
Other search	Frequency of selection	16%	11%	15%	7%	12%
Gmail	Frequency of selection	60%	66%	53%	52%	49%
Email services other than Gmail	Frequency of selection	43%	38%	50%	39%	41%
Facebook	Frequency of selection	61%	70%	61%	67%	68%
Twitter	Frequency of selection	22%	35%	18%	20%	17%
Google Maps, including Google WAZE	Frequency of selection	62%	50%	63%	53%	51%
Other mapping programs	Frequency of selection	11%	7%	13%	5%	9%
Other apps on an Android device	Frequency of selection	33%	23%	31%	26%	24%
Other apps on an iPhone or iPad	Frequency of selection	18%	20%	26%	18%	21%
Average Google use	Search+Gmail+Maps / 3	70%	67%	68%	64%	63%
Average Google alternative use	Sum of alternatives to Google / 3	23%	19%	26%	17%	21%
Technological sophistication	1 Low, 3 High	1,27	1,60	1,35	1,32	1,31
Experience online privacy violation	1=yes; 2=no	1,82	1,80	1,82	1,85	1,89
GDPR waiver Google	1=Yes, 2=I'm not sure, 3=No	1,77	1,39	2,44	1,87	2,02
Awareness	- 24 (not aware) to 24 (completely aware)	7,85	6,02	6,89	0,27	0,65
Consent	- 24 (no consent) to 24 (max. consent)	-18,38	5,89	-19,13	-12,86	-16,86
GDPR waiver for startups	1=Yes, 2=I'm not sure, 3=No	1,83	1,86	3,00	2,39	2,35
Comfortable using an EU alternative	1=Yes, 2=I'm not sure, 3=No	1,87	1,97	1,91	2,55	1,94
Usage of safe alternative	1=Certainly Not, 5=Certainly	4,08	3,84	4,39	2,18	4,00
Paying for safe alternative	1=Certainly Not, 5=Certainly	3,28	3,66	3,24	1,92	3,58
Regulatory satisfaction	-3 (dissatisfaction) to 3 (satisfaction)	-2,29	1,65	-2,39	-0,40	1,87
Education not yet completed	Selection	4%	6%	4%	6%	10%
Secondary education completed	Selection	30%	34%	26%	35%	34%
Some university or vocational diploma	Selection	12%	10%	10%	12%	8%
Vocational certification completed	Selection	16%	13%	15%	18%	16%
University degree completed	Selection	35%	31%	42%	25%	27%
Doctorate, post-doctorate	Selection	3%	3%	2%	2%	2%
Prefer not to answer	Selection	1%	3%	1%	1%	2%
UK	Selection	18%	22%	19%	23%	14%
USA	Selection	22%	21%	21%	18%	14%
France	Selection	19%	24%	17%	17%	24%
Denmark	Selection	20%	15%	24%	22%	23%
Germany	Selection	21%	18%	19%	20%	25%

Cluster 3 indicates similar characteristics to cluster 1. It has the second-highest value regarding the use of alternatives to Google and the highest value for not giving a GDPR waiver to startups and for regulatory dissatisfaction. Cluster 3 is rather aware of Google's practices and capabilities and does not approve of them. However, while cluster 3 is also willing to use a safe alternative, it is unsure about paying for it and whether to feel more comfortable if the alternative stems from Europe.

Cluster 4 differs from the others in two important ways. One, it indicates no willingness to use a safe alternative, and two, it is also not willing to pay for a safe alternative in preference to Google. It is the only cluster that demonstrates these characteristics. Interestingly, the fourth cluster strongly disapproves of Google's practices and capabilities. In addition, this cluster shows a consistent tendency towards the middle, indicating around 50/50 percent of awareness of Google's practices and capabilities, undecidedness towards providing a GDPR waiver for startups, and regulatory satisfaction. Cluster 4 has the lowest value on the use of alternatives to Google, but similar scores on the use of Google services compared with the other clusters.

Cluster 5 is unsure about giving startups a GDPR waiver, unsure about a European alternative, and indicates a 50/50 percent awareness of Google's practices and capabilities. Interestingly, the cluster is satisfied with regulators and does not approve of Google's practices and capabilities. Instead, it is willing to use a safe alternative and somewhat willing to pay for it. The usage of alternatives to Google is low, whereas the usage of Google's services is high.

Discussion

All in all, four of the five clusters indicate that they disapprove of Google's practices and capabilities. All are at least aware of 50% of the 24 listed practices and capabilities. However, extremely high scores on awareness are missing in clusters, indicating the lack of transparency and information. Four out of five clusters are willing to use a safe alternative, but only two are somewhat willing to pay for it. Thus, indicating the lack of understanding of digital and, more specifically, ad-based business models. And finally, we find evidence for Google's monopolistic position since all clusters demonstrate high usage of Google services and low usage of alternative services.

Privacy Paradox

Clusters 1, 3, 4, and 5 indicate the contradiction between disapproving of Google's practices and capabilities, but still using Google's services more frequently than alternative services. This contradiction confirms the privacy paradox in such a way that attitudes towards Google do not align with the actual behavior. Previous work argues that information asymmetries and incomplete information (Acquisti and Grossklags 2005; Buck et al. 2014), as well as the privacy calculus (Dinev et al. 2016; Xu et al. 2011) contribute to the explanation of the privacy paradox. We confirm both perspectives and also propose a complementary perspective. Hence, some consumers lack the information required for their privacy calculus (cluster 4 and 5), whereas others cannot locate a viable alternative (cluster 1 and 3). We propose that the lack of a valuable alternative provides a complementary explanation for the privacy paradox, especially in near-monopolistic markets. That is, cluster 1 and 3 disapprove of Google's practices and are willing to use a safe alternative but do not actually use it since it may not represent a viable alternative. Hence, consumers are behaving paradoxically, not because they are insufficiently informed or less aware of the privacy risks; contrarily, consumers are indeed informed and aware. However, they may continue using Google because they either (1) lack a valuable alternative, (2) lack access to alternatives because of Google's control over Android, Search, and Chrome, or (3) are unwilling to pay for alternatives.

Feasibility of new entry and possible regulatory response

Clusters 1 and 3 strongly value privacy as they are at least not sure whether they would provide Google or startups a GDPR waiver and are willing to use safe alternatives. In addition to consumer preference for privacy, new regulatory changes, such as the launch of GDPR, provide a fertile ground for easily entering the market according to the theory of NVM. However, the clusters indicate uncertainty about paying for safe alternatives, rendering it unattractive to enter. Hence, entry might be easier, but not attractive to attack, suggesting that consumers need to be better informed. Moreover, as conversions of

weak personalized ads perform significantly worse compared with highly personalized ads by Google, advertisers have a strong incentive to stay with Google. Thus, further rendering new market entry unattractive. Regulators could intervene in this context by better informing consumers about digital business models, monetization of personal data, and the logic and risks of personalized advertising. Understanding that a free and safe alternative might not be competitive to Google can help consumers change their minds about payments, especially since they generally disapprove of Google's practices.

Cluster 2 differs in three important ways. First, it provides Google a GDPR waiver; second, it is somewhat willing to pay for a safe alternative; and third, it mainly approves of Google's practices. However, as the cluster is undecided about granting startups a GDPR waiver, new entry might not be easy, as the startup would need to compete on different terms. But the market might be attractive to attack as new entry can monetize its safe business model. At the same time, it will be difficult to defend for Google because it cannot easily imitate safe business models without limiting its advantage of using vast amounts of personal data for highly personalized ads. Since this cluster provides, on an aggregated level, informed consent for Google's practices and is willing to use and pay for safe alternatives, no regulatory intervention might be necessary.

Although cluster 4 also strongly values privacy and strongly disapproves of Google's practices, it is not willing to use and pay for a safer alternative. This renders market entry difficult and unattractive and thereby unfeasible. It suggests to better inform consumers about the hidden costs of apparently free services, especially as this cluster indicates not to be well informed about Google's practices, and the logic and advantages of safe alternatives.

Cluster 5 exhibits similar characteristics to clusters 1 and 3. However, cluster 5 is also somewhat willing to pay for a safe alternative. This indicates that market entry is easy and attractive, and thereby, less or no regulatory intervention might be necessary.

Conclusion

We are reporting on a set of consumer surveys to examine consumers' attitudes towards new business models and towards firms that monetize consumers' personal information. Surveys assessed consumers' knowledge of Google's current business practices and their approval or disapproval of those practices, independent of their belief that those practices currently incurred. This allowed us to determine implicit informed consent: individuals who both knew and approved of these practices. We also studied their willingness to switch to alternative safer services, that is, services that do not violate privacy or monetize private information. We studied their degree of regulatory satisfaction, including the belief that regulators were taking adequate steps to inform consumers about potential privacy risks online and to protect them from these risks.

We collected survey responses from 1693 individuals about Google from Denmark, France, Germany, the UK, and the USA. Our cluster analysis confirms the privacy paradox – although consumers disapprove of Google's practices, they still use it – in four out of five clusters but indicates a new explanation: the lack of alternatives. Moreover, we draw on the theory of NVM to discuss whether market entry is feasible. Our results indicate that market entry is feasible for cluster 5 (willing to use and somewhat willing to pay for safe alternative and disapproving of Google's practices), rather difficult for cluster 2 (approving Google's practices and willing to grant it a GDPR waiver but somewhat willing to pay for safe alternative), rather difficult for cluster 1 and 3 (willing to use but unsure about paying for safe alternative and disapproving Google's practices), and completely unfeasible for cluster 4 (unwilling to use and pay for a safe alternative although disapproving Google's practices). We discuss which form of regulatory intervention might be necessary for each cluster. We argue that no intervention might be necessary for clusters 5 and 2, whereas clusters 1 and 3 and cluster 4 need different interventions in form of being better informed through regulation. Cluster 1 and 3 need to be better informed about the logic and advantages of safe alternatives to cope with their disapproval of Google's practices and their undecidedness to pay for a safe alternative. Cluster 4 needs to be better informed in two areas. First, about the hidden costs of apparently free services as it indicates not to be well informed about Google's practices. Second, about the logic and advantages of safe alternatives to cope with its disapproval of Google's services and its unwillingness to use and pay for safe alternatives. In contrast,

cluster 5 is willing to use and somewhat willing to pay for safe alternatives and, therefore, might render regulatory intervention unnecessary. Since cluster 2 provides, on an aggregated level, informed consent for Google's practices and is willing to use and pay for a safe alternative, no regulatory intervention might be necessary for this cluster neither.

This paper makes two contributions. The first is *practical*: The EU is seeking to develop new regulatory policies with multiple objectives. These include providing adequate (privacy) protections for consumers and providing a fair and level playing field that protects competition and limits the abuses created by firms that monetize personal information. Developing new regulatory policies entails a complex set of tradeoffs since regulatory policy must be informed by consumers' preferences, and if regulatory policy contravenes consumer preferences, it may be necessary for regulators to explain the hidden sources of harm to consumers. This study provides an input to the EU regulatory process. The second is *theoretical*: The privacy paradox indicates that consumers' online behavior often contradicts their expressed preferences for safe, secure, and private online interactions. We observed that while consumers express a preference for safer alternatives, they may be (a) unaware of the hidden costs of using existing Google offerings, (b) unable to find potentially safer alternatives with all the functionality they now obtain from Google, or (c) unwilling to pay for safer alternatives. This last point suggests that while consumers may have a real preference for privacy, the value placed on privacy may be quite low.

Our research has three limitations. First, we measure *expressed preferences*, what users claim to value, as opposed to what consumers actually value as revealed by their actions. This can introduce a significant level of inaccuracy, which must be considered before considering regulatory action. Second, because of the nature of the survey, we are exploring *what* users believe, but not *why* they believe it. Thus, while the results may be useful in explaining to regulators the situation they are facing, it offers little or no guidance on how to change this situation by altering what individuals know or believe. It is not a study of individual behavior but aggregate national beliefs. Third, the results reflect consumer attitudes towards Google and thereby limit the generalizable of our findings.

We propose three avenues for future research. First, examining informed consent for different empirical cases and to investigate which clusters emerge and which forms of regulation might be necessary in these cases. These investigations help to better understand and generalize consumer attitudes toward firms that monetize personal information and to theorize about appropriate forms of regulatory intervention. Second, when regulating in form of informing consumers, future research can shed more light on the effect of information on actual behavior. Research suggests that information alone does not always change behavior. Therefore, future work can explore which further forms of regulation are necessary to change consumer behavior and to provide a level playing field in among firms that monetize personal information. Third, moving beyond what users believe towards why they believe it. Hence, identifying the reasons why consumers didn't discontinue using the services given that they have expressed privacy concerns and willingness to use or pay for a safe alternative.

References

Acquisti, A., and Grossklags, J. 2005. "Privacy and Rationality in Individual Decision Making," *IEEE security & privacy* (3:1), pp. 26-33.

Acquisti, A., John, L. K., and Loewenstein, G. 2012. "The Impact of Relative Standards on the Propensity to Disclose," *Journal of Marketing Research* (49:2), pp. 160-174.

Adjerid, I., Peer, E., and Acquisti, A. 2018. "Beyond the Privacy Paradox: Objective Versus Relative Risk in Privacy Decision Making," *MIS Quarterly* (42:2), pp. 465-488.

Baumol, W. J. 1982. "Contestable Markets: An Uprising in the Theory of Industry Structure," *The American Economic Review* (72:1), pp. 1-15.

Buck, C., Horbel, C., Germelmann, C. C., and Eymann, T. 2014. "The Unconscious App Consumer: Discovering and Comparing the Information-Seeking Patterns among Mobile Application Consumers," *Proceedings of the European Conference on Information Systems (ECIS) 2014*, Tel Aviv, Israel.

Chakraborty, R., Vishik, C., and Rao, H. R. 2013. "Privacy Preserving Actions of Older Adults on Social Media: Exploring the Behavior of Opting out of Information Sharing," *Decision Support Systems* (55:4), pp. 948-956.

Clemons, E. K. 2018. *New Patterns of Power and Profit: A Strategist's Guide to Competitive Advantage in the Age of Digital Transformation*. Switzerland: Palgrave Macmillan.

Clemons, E. K., Gu, B., and Lang, K. R. 2002. "Newly Vulnerable Markets in an Age of Pure Information Products: An Analysis of Online Music and Online News," *Journal of Management Information Systems* (19:3), pp. 17-41.

Clemons, E. K., Wilson, J., and Jin, F. 2014. "Investigations into Consumers Preferences Concerning Privacy: An Initial Step Towards the Development of Modern and Consistent Privacy Protections around the Globe," *47th Hawaii International Conference on System Sciences*: IEEE, pp. 4083-4092.

Clemons, E. K., and Wilson, J. S. 2015. "Family Preferences Concerning Online Privacy, Data Mining, and Targeted Ads: Regulatory Implications," *Journal of Management Information Systems* (32:2), pp. 40-70.

Dinev, T., Albano, V., Xu, H., D'Atri, A., and Hart, P. 2016. "Individuals' Attitudes Towards Electronic Health Records: A Privacy Calculus Perspective," in *Advances in Healthcare Informatics and Analytics*. Springer, pp. 19-50.

Dinev, T., Bellotto, M., Hart, P., Russo, V., Serra, I., and Colautti, C. 2006. "Privacy Calculus Model in E-Commerce–a Study of Italy and the United States," *European Journal of Information Systems* (15:4), pp. 389-402.

Dinev, T., and Hart, P. 2006. "An Extended Privacy Calculus Model for E-Commerce Transactions," *Information Systems Research* (17:1), pp. 61-80.

Edelman, B. 2015. "Does Google Leverage Market Power through Tying and Bundling?," *Journal of Competition Law & Economics* (11:2), pp. 365-400.

Edelman, B., and Geradin, D. 2016. "Android and Competition Law: Exploring and Assessing Google's Practices in Mobile," *European Competition Journal* (12:2-3), pp. 159-194.

European Commission. 2017. "Antitrust: Commission Fines Google €2.42 Billion for Abusing Dominance as Search Engine by Giving Illegal Advantage to Own Comparison Shopping Service." Retrieved 01.05., 2019, from http://europa.eu/rapid/press-release_IP-17-1784_en.htm

European Commission. 2018. "Antitrust: Commission Fines Google €4.34 Billion for Illegal Practices Regarding Android Mobile Devices to Strengthen Dominance of Google's Search Engine." Retrieved 01.05., 2019, from http://europa.eu/rapid/press-release_IP-18-4581_en.htm

European Commission. 2019. "Antitrust: Commission Opens Investigation into Possible Anti-Competitive Conduct of Amazon." Retrieved 23.10., 2019, from https://europa.eu/rapid/press-release_IP-19-4291_en.htm

European Parliament. 2019. "Official Complaint by Spotify against Apple for Discrimination and Apple Music's Unfair Advantage over Spotify: The Power of Digital Platforms." Retrieved 19.02., 2020, from https://www.europarl.europa.eu/doceo/document/E-9-2019-002996_EN.html

European Parliament and the Council of the EU. 2016. "General Data Protection Regulation." Retrieved 17.01, 2020, from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN

FTC. 2019. "Google and Youtube Will Pay Record $170 Million for Alleged Violations of Children's Privacy Law." from https://www.ftc.gov/news-events/press-releases/2019/09/google-youtube-will-pay-record-170-million-alleged-violations

Gerber, N., Gerber, P., and Volkamer, M. 2018. "Explaining the Privacy Paradox: A Systematic Review of Literature Investigating Privacy Attitude and Behavior," *Computers & security* (77), pp. 226-261.

Iacobucci, E., and Ducci, F. 2019. "The Google Search Case in Europe: Tying and the Single Monopoly Profit Theorem in Two-Sided Markets," *European Journal of Law and Economics* (47:1), pp. 15-42.

Kokolakis, S. 2017. "Privacy Attitudes and Privacy Behaviour: A Review of Current Research on the Privacy Paradox Phenomenon," *Computers & security* (64), pp. 122-134.

Luger, E., Moran, S., and Rodden, T. 2013. "Consent for All: Revealing the Hidden Complexity of Terms and Conditions," *Proceedings of the SIGCHI conference on Human factors in computing systems*: ACM, pp. 2687-2696.

Phlips, L. 1983. *The Economics of Price Discrimination*. New York, USA: Cambridge University Press.

Porter, J. 2018. "Google Accused of Gdpr Privacy Violations by Seven Countries." *The Verge* Retrieved 1st May, 2019, from https://www.theverge.com/2018/11/27/18114111/google-location-tracking-gdpr-challenge-european-deceptive

Schenker, Y., Fernandez, A., Sudore, R., and Schillinger, D. 2011. "Interventions to Improve Patient Comprehension in Informed Consent for Medical and Surgical Procedures: A Systematic Review," *Medical Decision Making* (31:1), pp. 151-173.

Shiller, B. R. 2014. *First Degree Price Discrimination Using Big Data*. Brandeis Univ., Department of Economics.

Simon, H. A. 1997. *Models of Bounded Rationality: Empirically Grounded Economic Reason*. London, England: MIT press.

Weking, J., Hein, A., Böhm, M., and Krcmar, H. 2018. "A Hierarchical Taxonomy of Business Model Patterns," *Electronic Markets* (In Press), pp. 1-22.

Xu, H., Luo, X. R., Carroll, J. M., and Rosson, M. B. 2011. "The Personalization Privacy Paradox: An Exploratory Study of Decision Making Process for Location-Aware Marketing," *Decision support systems* (51:1), pp. 42-52.

Zhu, F., and Liu, Q. 2018. "Competing with Complementors: An Empirical Look at Amazon. Com," *Strategic Management Journal* (39:10), pp. 2618-2642.

Zuboff, S. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London, UK: Profile Books.

Appendix

Table 3: Factor Loadings Part 1

ID	Consent	Awareness	Regulatory Satisfaction
Q48	**0,793**	0,040	0,076
Q36	**0,791**	-0,048	0,120
Q26	**0,787**	0,049	-0,019
Q38	**0,778**	0,024	0,001
Q24	**0,772**	-0,051	0,164
Q28	**0,769**	-0,039	0,133
Q22	**0,761**	-0,036	0,113
Q46	**0,755**	0,040	0,017
Q40	**0,755**	-0,083	0,108
Q14	**0,753**	-0,053	0,168
Q12	**0,752**	-0,047	0,151
Q44	**0,752**	0,025	0,035
Q34	**0,750**	0,022	0,018
Q18	**0,746**	0,000	0,131
Q30	**0,742**	-0,008	0,044
Q42	**0,739**	-0,059	0,171
Q16	**0,733**	-0,058	0,176
Q50	**0,728**	-0,033	0,195
Q52	**0,721**	-0,044	0,159
Q32	**0,716**	0,023	-0,018
Q20	**0,708**	0,134	-0,046
Q10	**0,686**	0,117	-0,036
Q59	**0,656**	0,120	-0,014
Q57	**0,635**	0,124	-0,055
Q31	-0,003	**0,757**	-0,021
Q29	0,020	**0,746**	-0,007
Q33	0,003	**0,745**	0,075

Table 4: Factor Loadings Part 2[2]

ID	Consent	Awareness	Regulatory Satisfaction
Q27	-0,009	**0,741**	-0,062
Q39	0,034	**0,735**	-0,026
Q41	0,005	**0,730**	-0,104
Q35	0,000	**0,698**	-0,024
Q21	-0,022	**0,692**	0,036
Q23	-0,055	**0,676**	-0,110
Q13	-0,026	**0,660**	-0,032
Q51	0,051	**0,658**	0,088
Q11	-0,020	**0,651**	-0,033
Q15	-0,020	**0,648**	-0,044
Q37	-0,009	**0,624**	0,148
Q45	0,022	**0,607**	0,142
Q43	0,016	**0,607**	0,167
Q49	0,033	**0,586**	-0,013
Q17	0,005	**0,576**	0,046
Q25	0,030	**0,553**	0,244
Q47	0,011	**0,548**	0,189
Q58	0,009	**0,461**	0,309
Q19	0,020	**0,383**	0,273
Q56	0,039	**0,379**	0,359
Q9	-0,021	**0,329**	0,273
Q67	0,324	0,060	**0,679**
Q68	0,324	0,082	**0,652**
Q66	0,311	0,087	**0,636**

[2] Q66 = Do you believe that regulators do an adequate job ensuring that you are aware of the way your online data is used and reused by your service providers? Q67 = Do you believe that regulators do an adequate job ensuring that you are aware of the way you may be harmed by the use of your online data by your service providers? Q68 = Do you believe that regulators do an adequate job ensuring that you cannot be harmed by the use of your online data by your service providers?