Why Does No One Read Website Privacy Policies[1]?

Grazia Cecere, Telecom Ecole de Management, Institut Mines Telecom, RITM-University of Paris Sud and Digital Society Institute, France. Email: grazia.cecere@telecom-em.eu

Fabrice Le Guel, RITM-University of Paris Sud, France. Email: fabrice.le-guel@u-psud.fr

Nessrine Omrani, PSB-Paris School of Business & Chair D^3, France. Email: n.omrani@psbedu.paris

[1] Work supported by the DAPCODS/IOTics ANR 2016 project (ANR-16-CE25-0015).

Why Does No One Read Website Privacy Policies?

Abstract:

Online privacy policies should decrease information asymmetry between Internet users and websites as they might inform individuals about how firms collect and use personal data. However, the number of individuals that read privacy policies is very limited according to recent studies based on non-declarative responses and clickstream data. Few published studies have investigated why people do not read privacy policies. In this article, we focus on the reasons behind the likelihood of *not reading or only partially reading* privacy policies. For this purpose, we analyze a large sample including individuals' details from 28 European countries in 2015 with a sample of 13,783 respondents. We show that a potential group of citizens should be better informed if privacy policies are improving.

Keywords: Privacy policies, minority informed, privacy paradox.

1. INTRODUCTION

Privacy policies, also named privacy statements, privacy notices or 'terms of use', are an example of well-known 'fine print' in the literature in law and economics (Bakos et al. 2014). Applied to the Internet or mobile applications, online privacy policies, should

detail how websites collect, store, and use personal data. More precisely, privacy law in European countries (as in the U.S.) is based on the free market 'notice-and-choice' principle (Schwartz and Solove, 2009; Cranor, 2012; Zimmeck et al., 2014; Athey et al., 2017) which implies that once individuals use the website or a given mobile application, they implicitly accept the content of the privacy policy. There is no legal obligation requiring website publishers to display formatted content for privacy policies, just some recommendations like the United States Federal Trade Commission's fair information practice which is similar to the European Union Data Protection Directive principle recently upgraded to the General Data Protection Regulation (GDPR).

From an economic viewpoint, privacy policies are important as they can decrease information asymmetry between Internet users and websites regarding how personal data are gathered and used. For example, Eastlick et al., (2006) and, Smith et al. (2011), showed that privacy policies may vehicle an individual's trust. While clearly and compactly displaying, privacy policy information to consumers has a positive impact on purchases from online retailers (Tsai and al. 2011). Bowie and Jamal (2006) also found that 'trustworthy' firms may have a competitive advantage because consumers are more willing to disclose personal details (Schoenbachler and Gordon, 2002). Ultimately, information asymmetry is likely to become even more contentious in the future with the Internet of things.

More importatly, privacy policies would have been proposed by the US FTC to encourage a marked-driven approach to privacy "as a means of facilitating competition

over privacy practices." (Beales, 2002)[2]. Privacy policies may also be considered as a loud signal to consumers about the trustworthiness of the website (Martin, 2016). Privacy policies are important because in an environment where the regulator is unable to verify that each website is respectful of privacy, privacy policies can be considered a first step to complement the existing regulation with a more efficient self-regulation.

In fact, "there are several specific market mechanisms by which privacy might be regulated, rather than by the government. The primary one is contract" (Ben-Shahar and Strahilevitz, 2016, p. 3). However, privacy policies will form the basis for the self-regulatory notice only if what is declared is really done by the websites and more broadly by firms collecting personal details and only if consumers take into account (i.e. read and understand) privacy policies. The literature in law and economics about the 'fine print' showed that if only a minority of consumers read the fine print, this 'informed minority' would be sufficient to encourage most actors to write better notices respecting their commitments. We could also think that if a sufficient number (to achieve the 'informed minority') of web users read privacy policies, a competition over privacy-respectful practices could arise, as a privacy policy becomes a means of differentiation for websites (which was the assumption of the FTC at the origin of the 'notice-and-choice' principle, Beales, 2002).

[2] In a 2002 speech, Howard Beales, Director of the FTC Bureau of Consumer Protection, said, "First, privacy notices should be viewed as a means of facilitating competition over privacy practices. Their goal should be to help consumers understand what information is collected about them and what is done with that information, not to simply scare consumers into opting out of information sharing." In Beales, H. J., (2002), Privacy Notices and The Federal Trade Commission's 2002 Privacy Agenda, part "The Role of Privacy Notices."

However, the literature showed that very few individuals read (and still less understand) privacy policies because they are too long and there are too difficult to understand (Cranor, 2012; Strahilevitz and Kugler, 2016). Then, as websites know that nobody (or very few, meaning less than the so-called 'informed minority') reads – including supervisory authorities - privacy policies, there are no incentives for firms to declare what they really do with personal data collected. Privacy policies become a poor quality-signal in a market with information asymmetry (Akerlof, 1970: Vila et al., 2003; Reay et al. 2009). An adverse selection mechanism could even take place (Edelman, 2011) where only 'poor quality' privacy policies (i.e. websites and more broadly firms non-respectful of privacy) should be the norm corresponding to pooling Nash equilibria (Reay et al., 2009), a major tenet of signaling theory, which, *a priori*, describes the current situation. To switch to a 'separating equilibrium' where web users can differentiate high versus poor quality privacy policies (then separate firms non-respectful of privacy from respectful ones), it seems important to better understand why a significant majority of web users do not read privacy policies and how this situation can be improved.

To reach this goal, we use an original Eurobarometer database with 13,783 representative individuals from 28 European countries collected in 2015. Individuals declare if they read and understand the privacy statement of the websites visited and the reasons of not reading, or partially reading, privacy policies. To our knowledge, no empirical multi-country works have studied the factors explaining why people do not read privacy policies. We exploit rich information (unavailable with clickstream data) to identify factors affecting whether people read privacy policies and the characteristics of the individuals related to each reason. For this purpose, we estimate a cluster analysis

to identify a clear group of individuals. We principally show that a potential group of citizens should be better informed if privacy policies are improving. Another important finding of our paper is about how the surveys should question individuals about their consideration of privacy policies and more broadly about privacy enhancing technologies. It seems that the declared responses in the case of privacy are of poor quality because, as stipulated by Acquisti et al (2015), when studying individual choices facing privacy issues, we are in an environment filled with paradoxes (called the 'privacy paradox'). Tucker (2014) also calls this phenomenon the "Internet users' perceptions of control": when people have the illusion of control over the collection and use of their personal data, they are more willing to entrust the company. With privacy policies, most individuals declare read and understand these notices - while it is possible that very few read them - because they may feel they have control over their personal information.

The article is organized as follows. Section 2 provides a review of the literature dealing with privacy policies in different theoretical frameworks and presents the main hypothesis tested. Section 3 presents the data used and the methodology, section 4 presents and discusses the results, and section 5 provides a conclusion.

2. RELATED LITERATURE

Our paper contributes to two literature streams: the economics of privacy and the economy of 'fine print'.

2.1. How Information Asymmetry between data holder and data subject

According to Acquisti (2013), "the economics of privacy attempts to study the costs and benefits associated with the protection or disclosure of personal data - for the data subject, the data holder, and for society as a whole." As the value of personal data is difficult to assess because it is an information good and usually overestimated by individuals; offer and demand do not match, and we cannot consider that a strict market for personal details exists according to the definition of Laudon (1996) or Varian (1996). However, even if the relevant market for personal data is nowadays difficult to determine , the literature agrees on the fact that there is a problem of asymmetric information in this type of market (Akerlof, 1970). Indeed, applied to the Internet (but this is also the case for all connected devices), information asymmetry occurs between the web user and the website visited. In such an environment, information asymmetry is on both sides of the market but with a large disadvantage for consumers. When purchasing online, for example (or when visiting a given website), an online merchant may have limited information about the consumer's tastes, reservation price, identity, and so on (Acquisti and Varian, 2005). However, after the purchase (or after the visit), the online merchant has been able to gather different types of personal data, and personal information can be posted voluntarily, legally observed (e.g., Internet browsing preferences and location data), or inferred, based on the analysis of personal

details (OECD, 2013). A consumer may not know how the merchant will use the personal information revealed as part of the transaction (Acquisti and Grossklags, 2005). Additionally, some personal data may be used for discrimination (Datta et al., 2015) or sent to third parties, such as data aggregators, data brokers or ad companies (Tucker, 2014) without the consumer's knowledge (Varian, 1996). Akçura and Srinivasan (2005) showed that this 'second usage' of personal data may negatively impact consumers' welfare, and 'first usage' helps firms better interact with customers providing their personal data.

This lack of information about how personal data are used by firms affects individual behavior in different ways according to the level of risk aversion. For one, consumers may perceive greater risk and uncertainty when dealing with merchants. As a result, they may be less willing to complete transactions with those merchants. If there is incomplete information, consumers are not even aware that their personal information could be exchanged or misused, and it may make them more likely to engage in such risky transactions. To minimize these risks and decrease information asymmetry, several solutions have been proposed. First, regulatory frameworks could protect consumers. For example, European countries are regulated by the data protection directive adopted in 1995 (Directive 95/46/EC) and recently up dated. However, Cecere et al. (2015) showed that the law was not sufficient because it does not account for highly heterogeneous privacy preferences.

Secondly, self-regulation (with, for example, the privacy-by-design framework or by using online 'trust' authorities) could be a better solution. However, Edelman (2011) showed that only low-quality websites (in terms of online privacy and safety) seek and receive certification, which leads to adverse selection; it seems that self-regulation does not work when some agents (here, the websites) have hidden types.

A third type of tool used to decrease information asymmetry should be the Privacy Enhancing Technologies (PET), but few people use it because it requires technical skills (see, for example, the Platform for Privacy Preferences project[3] - P3P, Beaty et al., 2007).

Finally, one of the best-known ways to inform consumers could be the privacy policy. A privacy policy is a statement that discloses some or all the ways a website manages personal data. In such a scheme, we can consider a website a credence good because the user cannot verify how personal data are used by the website. Then, we can suppose that a privacy policy allows converting a credence good to an experience good (McDonald and Cranor, 2008)[4]. In this context, privacy policy can be considered a signal (Vila et al., 2003), and reading privacy policies could reduce information asymmetry.

[3] The Platform for Privacy Preferences project (P3P) aimed to develop a user agent allowing websites to declare their intended use of information collected. This project was one of the most famous technologies to inform consumers. However, stakeholders (i.e. websites, web browsers, and Internet users) have not widely adopted this technology.

[4] A search good is a service or product where the characteristics can be determined before it is consumed (i.e. *ex ante*). An experience good is a product or service where characteristics can only be ascertained upon consumption (i.e. *ex post*). A credence good (or post-experience good) is a service or product for which the consumer cannot ascertain the characteristics either *ex ante* or *ex post*.

2.2. Privacy Policy: Theoretical Background

A privacy policy can also be considered standard form contracts - called 'fine print' in the literature in law and economics (Bakos and al, 2014). This type of contract is used for billions of commercial transactions where a basic contract form is printed (a 'boilerplate') showing guaranties and how to make a claim. A first strand of literature justifies the non-intervention of the regulator based on the so called 'informed-minority hypothesis' by arguing that the self-regulation of sellers works if a minority of consumers are aware of the contract terms. Specifically, Schwartz and Wilde (1979) assume that consumers have different levels of ability to read and understand contract terms, and they show that if a sufficient number of consumers are informed about pricing and contract terms, at the equilibrium, sellers will provide efficient contract terms and low prices. As sellers cannot, at a low cost, discriminate between informed and uninformed buyers, sellers choose to provide an optimal fine print after performing a cost-benefit analysis.

The 'informed-minority hypothesis' is conditioned by the fact that a significant minority of individuals *should* read privacy policies. Previous empirical works have investigated the rate of reading privacy policies, but the empirical results are inconsistent. Indeed, reading privacy policies implies support for the 'cost' associated with the time spent reading them, which might represent a barrier. From a cognitive viewpoint, privacy policies do not have a standardized format, and the literature shows that privacy policies are usually too long to be read and too difficult to be understood by web users. This implies that rational actors might consider the trade-off between the benefit of being informed and the time spent reading the privacy policy. McDonald and Cranor

(2008) measure the opportunity cost for a consumer to read privacy policies. They consider the time spent reading them as a cost. They use a list of the 75 most popular websites and assume an average reading rate of 250 words per minute to find an average reading time of 10 minutes per policy. They estimate the value of time as 25% of an average hourly salary for leisure and twice the wages for time at work. They present a range of values and found the national opportunity cost for just the time spent reading policies is on the order of $781 billion. They also determined that "if every U.S. Web user read the privacy policy at every site visited, the time spent reading privacy policies would total an estimated 44.3 billion hours per year." Privacy policies are also often written in ways that are truly confusing (Reidenberg et al., 2016), and they also change over time since firms include a modification clause in the notice which allows them to make such changes (Ben-Shahar and Strahilevitz, 2016). The complexity of contract terms dealing with privacy is also a major obstacle for efficient private contracting (Ben-Shahar and Strahilevitz, 2016).

Although the literature mainly indicates that actual privacy policies are inefficient and difficult to read and understand, most surveys report high rates among respondents declaring that they read and understand the privacy policy when they visit a website. For example, Milne and Culnan (2004) found that about 83.7% of respondents declared to be readers[5]. The article of Jensen et al. (2005)[6] showed that 43 % of individuals are

[5] The study used an online survey of 2,468 U.S. adult Internet users. Respondents were asked how frequently they read privacy policies using a scale ranging from 1 (never read) to 5 (always read). Among the readers, 4.5% indicated they always read, 14.1 % declared they frequently read, 31.8 % read sometimes, and 33.3 % indicated they rarely read.

[6] The study was an online survey with 175 volunteers recruited through email and advertisements on academic websites.

likely to read the privacy policy of an ecommerce site before buying anything. In the Global Internet Survey (2012)[7], 72% of respondents reported that they read privacy policies sometimes or most of the time, and 16% always read the privacy policy. The recent study of Turow et al. (2015) showed that half of online Americans know what a privacy policy is[8]. Strahilevitz and Kugler (2016) conduct a study through two experiments in which consensus-weighted samples of more than a thousand Americans read short excerpts from Facebook, Yahoo, and Google's privacy policies concerning the use of facial recognition software and automated content analysis in email. The study shows that courts and laypeople may understand the same privacy policy language quite differently. According to Strahilevitz and Kugler (2016), even when consumers do read privacy policies, their beliefs about the nature of their bargains with technology companies seem to depend more on their pre-existing expectations than on the terms of the policies.

Conversely, some studies using Internet clickstream data (i.e. non-self-reported behavior) showed that reading rates of privacy policies was very low (Tsai and al. 2011). For example, Bakos et al. (2014) study individual behavior by tracking the Internet browsing of 48,154 monthly visitors to 90 online software companies in January 2007, and the authors shows that only about 0.22% of individuals read end-user license agreements (EULAs) of the websites they visited. Using individual data usages in a lab experiment, Jensen et al. (2005) showed that only 25.9% of volunteers consulted privacy

[7] About 10,000 individuals in 20 countries were asked about their attitudes towards the Internet and behaviors online.
[8] The analysis was based on an online survey conducted September 12-18, 2014 among a sample of 1,066 adult Internet users. See also Pew Research Center, November 2014, "What Internet Users Know About Technology and the Web."

policies, compared to the 43% declared in the survey. The difference between reported behavior and actual behavior might be inflated because "subjects knew they were being observed, and what the purpose of the experiment was. They therefore likely took more care and were more thorough in their decision-making process than they normally would." (ibid, p. 215). In a recent experiment conducted in London by security firm F-Secure (2014), researchers built a portable Wi-Fi access point on a prominent street in London. After thirty minutes, 250 devices connected to the Wi-Fi access, and 33 people used the hotspot for web searches or to send emails despite the presence of a Terms and Conditions page asking hotspot users to "give their firstborn child or a beloved pet" in return for free access[9].

2.3. 'Privacy –Policy- Paradox'

These differences in the empirical literature could be the result of response bias or a sign of the so called 'privacy paradox' (Acquisiti and Grossklags, 2005; Acquisti and Gross, 2006; Strahilevitz and Kugler, 2016). This literature stream advocates that the discrepancies between actual behavior and stated preferences are due to individuals not being able to assess their real privacy concern level (Acquisti and Gross, 2006), stating a rationality problem where what is declared by the consumer is different from what is actually observed. Behavioral economics also gives insights on the motivation that justifies why individuals do not read the privacy policy: because they believe that they

[9] Even the Chief Justice of the United States Supreme Court admitted he "doesn't usually read the computer jargon that is a condition of accessing websites", see "Supreme Court Chief Justice Admits He Doesn't Read Online EULAs Or Other 'Fine Print'", in *techdirt.com*, https://www.techdirt.com/blog/?tag=john+roberts. See also FTC, (2010), "Protecting consumer privacy in an era of rapid change. A proposed framework for businesses and policymakers", Preliminary FTC Staff Report, p. 27.

are fully informed about privacy issues (Vila et al., 2003; Acquisti and Grossklags, 2005) or because they the trust websites (Martin, 2016). Martin (2016) studied the impact of privacy policies on trust online using a factorial vignette study. Results show that invoking privacy policies decreases trust in a website. This may be due to the lack of understanding or reading privacy policies since, according to Martin (2016), they could be considered a complement for consumers' trust.

Then, on the Internet and applied to privacy policies, the 'minority informed' seem to be practically nonexistent, and the equilibrium better describes a situation where firms know that no one reads privacy policies. Privacy policies are just displayed 'in case if' the regulator becomes concerned. From a theoretical viewpoint, if the signal (i.e. the privacy policy) is not effective at reducing the information asymmetry, the equilibrium is a lemons market (Akerlof, 1970; Vila et al. 2003). In our case, the expectation regarding the quality of the privacy policy is driven downwards.

Therefore, understanding why people do not read privacy policies seems fundamental to hoping than one day the 'informed-minority hypothesis' will apply and initiate a 'virtuous circle' encouraging firms to respect privacy and consumers to read privacy policies. This is the aim of the next section.

3. DATA AND METHODOLOGY

To understand why individuals do not read privacy policies, we use an empirical study based on original data collected by the EU[10] which include 13783 observations from over

[10] We use the Eurobarometer Special Survey n°431.

28 EU countries. We rely on the literature review to identify the key variables affecting the reasons why people do not read privacy policies. One of the Eurobarometer survey questions asks individuals why they do not read or only partially read privacy statements (Table 1). The advantage of this study over previous work is that we have detailed information about the motivation for not reading privacy policies. To grasp the motivations for not reading privacy policies, the empirical section includes the hierarchical cluster analysis as it permits the uncovering of groups of individuals. We use the unweighted pair-group method, which is also called group average linkage. The distance between two clusters is computed as the average distance of all pairs of individuals. The cluster analysis is conducted on seven dummy variables about the motivation of not reading privacy policy (Table 1): not honored, protect law, findability, not important, too long, unclear, and policy present. Most individuals, or about 62%, declare that the privacy policies are too long to be read, and about 35% declare that privacy policies are unclear.

[Table 1 near here]

A hierarchical cluster analysis (by using the unweighted pair-group method) conducted on these seven dummy variables shows that some groups of reasons for not reading or partially reading privacy polices exist. Table 2 presents the variables belonging to each cluster. The results of the cluster show that four groups of motivations can justify why individuals do not read or partially read privacy policies. Cluster 1 includes the variables TOO LONG and UNCLEAR, which suggests that this cluster is associated with the

understanding of privacy policy. A second cluster is composed by the variables PRIVACY PRESENT and PROTECT LAW, which is clearly associated with trust of the institutions. Cluster 3 includes the variables NOT IMPORTANT and FINDABILITY, which could be associated with a lack of interest in the privacy policy and more broadly, a low preference for privacy. Cluster 4 only includes the variable NOT HONORED, which is associated with a lack of trust of websites.

[Table 2 near here]

Statistically, privacy policies are not read (or partially read) due to the length, weak understandability, or both, which is in accordance with the literature (Cranor, 2012). Then, we can suppose that clearer or shorter privacy policies could encourage this part of the population to read these statements (then increasing the size of the "minority informed" group). Conversely, individuals who declare that they do not read privacy policies because they are unimportant or because they cannot find them (probably due to a lack of interest) will not be susceptible to changing their behavior if privacy policies are made more readable. Similarly, better privacy policies should not substitute a lack of trust of websites. Finally, if privacy policies can be considered a means to manage the risks of disclosing personal information online, these are currently unreadable, and thus probably not read. A strategy for web users would be to trust in national public authorities, European institutions, or laws.

Individuals belonging to (i.e. reasons to not read or partially read privacy policies) clusters 2, 3, and 4 are not potential readers, even if privacy policies are improved. To

rely on institutions is not sufficient because these institutions also cannot read all privacy policies and check if what is declared is really done by websites. In such a scheme, only people who principally declare not reading (or partially reading) privacy policies could become the potential informed group to pressure websites to respect privacy if better solutions to manage the risks of disclosing personal information are proposed. What is the profile of this community? Is the number of potential adopters sufficient to inaugurate a virtuous dynamic if privacy policies become more readable as suggested by Bakos et al. (2014)?

Table 3 presents the descriptive statistics of the main variable of interest and the control variables of our model. To measure the level of asymmetry of information, the variable KN_AUTHORITY considers if individuals know the national regulatory authority of the privacy issues. The dummy variable TRUSTGOVERNMENT measures to what extent individuals' trust national authorities, and about 45.2% of individuals declared they trust the national institution. The variable MISUSE measures the individuals' privacy policy concerns; a value of 1 indicates low privacy concerns while value 4 is associated with high privacy concerns.

The demographic explanatory variables include AGE, GENDER, MARRIED, and DEGREE, which measure age, gender (value 1 if the individual is male and 0 otherwise), marital status (value 1 if the individual is married and 0 otherwise), and degree obtained (value 1 if the individual has a degree and 0 otherwise), respectively. The mean age is 44.57, and 45.61% of the respondents are male, while 53.54% are married. Job positions are

measured through dummy variables: STUDENT, SELFEMP, EMPLOYED, and NOWORK; 6.55% of the respondents are students, 6.90% are self-employed, 39.78% are employed, and 53.32% are retired. These variables can also be considered measures of income[11]. A set of mutually exclusive binary variables indicates whether the individual lives in a large town, middle-sized town, or rural area. The reference variable is RURAL (to our knowledge, no other studies on this topic consider these spatial variables). While 27.42% of the respondents live in a large town (LARGE), 41.66% live in a small or middle-sized town (MIDDLE), and 30.92% live in a rural area. The frequency of Internet use is measured by the variable FREQ_INT. A value from 1 to 6 is assigned according to the degree of Internet use, and 59.27% (which takes the value 6) use the Internet daily or almost daily.

[Table 3 near here]

[11] For this purpose, the job position is a good indicator of people's living conditions with respect to income. People's incomes reflect differences in living standards among European countries rather than the real living conditions of people in their countries.

To estimate the reasons why individuals in each cluster do not read privacy policies, we deeply analyze the descriptive statistics of each cluster to identify some important regularities. Cluster 2 is associated with a high percentage of individuals that know the a national authority of regulation exists, and they largely trust national authorities. Cluster 3 is characterized by a low percentage of individuals who know the national authorities for privacy regulation. Cluster 4 is characterized by a high level of privacy concern and reduces the asymmetry of information.

[Table 4 near here]

4. RESULTS AND DISCUSSION

To estimate the probability of belonging to a specific cluster, we run a set of dichotomous Logit estimations aimed at determining the variables that can influence the behavior of individuals according to their beliefs (clusters). Table 5 presents the results of the Logit estimations.

[Table 5 near here]

Individuals that declare privacy policies are too long to read and unclear belongs to Cluster 1. Being concerned with privacy issues increases the probability of this group of individuals not reading privacy policies. While individuals are concerned about privacy, they do not want to support the opportunity costs to read privacy policies.

Individuals who are not concerned or informed about privacy issues declare they do not read privacy policies because it is difficult to find them or because privacy policies are unimportant. Belonging to Cluster 3 is also associated with low education compared to the other clusters as the variable degree is negative and significant, and these individuals are not frequent users of Internet. Specifically, this group of individuals is less educated and informed about privacy issues which suggests that regulatory authorities have also dedicated attention to these individuals.

In Cluster 4 (the lack of confidence in websites) (NOT HONORED), the results show that being informed about privacy issues is measured with the variables KNOWLEDGE AUTHORITY and MISUSE and do not increase the probability of reading privacy policies; because of a lack of confidence in the website as these individuals believe that the privacy policy is not honored in any case. Then, the reduced level of asymmetry of information does not increase the probability of reading the privacy policy. These individuals do not also trust national institutions. The sociodemographic variable male and individuals with a degree declare they do not read privacy policies as they are not honored by the website in any case.

The decision to not read privacy policy once individuals trust the institution (PROTECTED LAW AND POLICYPRESENT), or individuals who belong to CLUSTER 2, is associated with trust in the national government (TRUSTGOV) and information related to national regulatory authority of privacy (KNOWLEDGE AUTHORITY). Specifically, individuals who trust institutions do not attribute importance to privacy policies. Improving the

readability of privacy policies should not increase the size of the informed minority group for individuals belonging to this profile. However, risk aversion measured with the variable MISUSE is associated with the probability of not reading privacy policies. Then, one can believe that risk adverse individuals are more likely interested in being informed about privacy issues.

5. CONCLUSION

Information privacy law in European countries (as in the U.S.) is based on the free market notice-and-choice principle (Cranor, 2012; Zimmeck et al., 2014) which implies that once individuals use the website they implicitly accept the content of the privacy policy. Despite reading rate differences, it is important to understand the motivation that leads individuals to read or not read privacy policies.

The literature on law and economics recognizes that most buyers do not read the fine print due to a lack of time or because these contracts are difficult to understand. This could lead to a market failure due to a problem of imperfect information justifying regulation: if nobody reads the contracts, sellers will not be motivated to differentiate from competitors by offering a more readable contract with better guarantees than the minimally-required legal protections. If a minority of individuals is *informed* (the so-called 'informed minority'), this might force firms to respect the term and conditions. These rules are also applicable in the context of Internet with privacy policies. We show that nobody read and understand privacy policies, but we also demonstrate that a potential population of readers exists if privacy policies become shorter and clearer. Institutions should encourage the development of norms for privacy policies to engage

in a virtuous circle where a minority of informed citizens will be sufficient to motivate Internet stakeholders to using personal information to do what they declare in the privacy policies. Then, these statements could become a tool for differentiation and competition as originally intended by the Federal Trade Commission.

However, when thinking about a better format for privacy policies, it is important to avoid the 'transparency paradox,' which means "for a privacy policy to be actually transparent, the policy needs to be detailed and point out exactly who interacts with the data, when, how and to what end. However, this detail renders the texts so complex that no one reads them, let alone understands them" (Arnold et al., 2015, p.30; See also Nissenbaum, 2011). This complexity opens fertile ground for the incorporation of behavioral factors into the understanding of privacy decision-making (Adjerid, Samat, and Acquisti, 2016).

REFERENCES

Acquisti, Alessandro, and Jens Grossklags. 2005. Privacy and Rationality in Individual Decision Making. *IEEE Security & Privacy Magazine*, 3(1):26–33.

Acquisti, Alessandro, and Ralph Gross. 2006. Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook. Pp. 36–58 in *Privacy Enhancing Technologies*, edited by George Danezis and Philippe Golle. New York: Springer.

Acquisti, Alessandro. 2013. The Economics of Privacy: Theoretical and Empirical Aspects. Working paper. Carnegie Mellon University, Pittsburgh, PA.

Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. 2015. Privacy and Human Behavior in the Age of Information. *Science*, 347(6221):509–515.

Acquisti, Alessandro, and Hal R. Varian. 2005. Conditioning Prices on Purchase History. *Marketing Science*, 24(3):365–381.

Adjerid, Idris, Sonam Samat, and Alessandro Acquisti. 2016. A Query-Theory Perspective of Privacy Decision Making. *Journal of Legal Studies*, 45:S97–S121.

Akçura, M. Tolga, and Kannan Srinivasan. 2005. Customer Intimacy and Cross-Selling Strategy. *Research Note: Management Science*, 51(6):1007–1012.

Akerlof, George. 1970. The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 84(4):488–500.

Arnold, Rene, Annette Hillebrand, and Martin Waldburger. 2015. *Personal Data and Privacy*. London: Ofcom.

Athey, Susan, Christian Catalini, and Catherine Tucker. 2017. The Digital Privacy Paradox: Small Money, Small Costs, Small Talk. *NBER Working Paper* No. 23488.

Bakos, Yannis, Florencia Marotta-Wurgler, and David R. Trossen. 2014. Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts. *Journal of Legal Studies*, 43(1):1–35.

Beales, H. J.. 2002. Privacy Notices and The Federal Trade Commission's 2002 Privacy Agenda, part "The Role of Privacy Notices."

Beatty, Patricia, Ian Reay, Scott Dick, and James Miller. 2007. P3P Adoption on E-Commerce Web Sites: A Survey and Analysis. *IEEE Internet Computing*, 11(12):65–71.

Ben-Shahar, Omri, and Lior Jacob Strahilevitz. 2016. Contracting Over Privacy: Introduction. *Journal of Legal Studies*, 45(52), S1–S11.

Bowie, Norman E., and Karim Jamal. 2006. Privacy Rights on the Internet. *Business Ethics Quarterly*, 16(3):323–342.

Cecere, Grazia, Fabrice Le Guel, and Nicolas Soulié. 2015. Perceived Internet Privacy Concerns on Social Networks in Europe. *Technological Forecasting and Social Change*, 96:277–287.

Cranor, Lorrie Faith. 2012. Necessary but Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice. *Journal on Telecommunications and High Technology Law*, 10(2):273–307.

Datta, Amit, Michael Carl Tschantz, and Amupam Datta. 2015. Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination. *Proceedings on Privacy Enhancing Technologies*, 1:92–112.

Eastlick, Mary Ann, Sherry L. Lotz, and Patricia Warrington. 2006. Understanding Online B-to-C Relationships: An Integrated Model of Privacy Concerns, Trust, and Commitment. *Journal of Business Research*, 59(8):877–886.

Edelman, Benjamin. 2011. Adverse Selection in Online 'Trust' Certifications and Search Results. *Electronic Commerce Research and Applications*, 10(1):17–25.

F-Secure. 2014. *Tainted Love: How Wi-Fi Betrays Us.* Available from https://fsecureconsumer.files.wordpress.com/2014/09/wi-fi_report_2014_f-secure.pdf

Federal Trade Commission [FTC]. 2009. *FTC staff report: Self-regulatory Principles for Online Behavioral Advertising*. Available from http://www.ftc.gov/os/2009/02/P085400behavadre port.pdf.

Federal Trade Commission [FTC]. 2000. Fair Information Practices in the Electronic Marketplace. Washington, DC: Federal Trade Commission.

Global Internet Survey. 2012. *Summary Report.* Available from http://www.internetsociety.org/sites/default/files/rep-GIUS2012global-201211-en.pdf

Jensen, Carlos, Colin Potts, and Christian Jensen. 2005. Privacy Practices of Internet Users: Self-reports Versus Observed Behavior. *International Journal of Human-Computer Studies*, 63:203–227.

Laudon, Kenneth C. 1996. Market and Privacy. *Communication of the ACM*, 39(9):92–104.

Martin, Kirsten. 2016. Do Privacy Notices Matter? Comparing the Impact of Violating Formal Privacy Notices and Informal Privacy Norms on Consumer Trust Online. *Journal of Legal Studies*, 45:191-214.

McDonald, Aleecia M., and Lorrie Faith Cranor. 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society*, 4(3), 543–568.

Milne, George R., and Mary J. Culnan, 2004. Strategies for Reducing Online Privacy Risks: Why Consumers Read (or Don't Read) Online Privacy Notices. *Journal of Interactive Marketing*, 18:15–29. doi: 10.1002/dir.20009

Nissenbaum, Helen. 2011. A Contextual Approach in Privacy Online. *Daedalus*, 140(4):32–48.

OECD. 2013. *Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value.* Paris: OECD Publishing.

Reay, I., S. Dick, and J. Miller. 2009. An Analysis of Privacy Signals on the World Wide Web: Past, Present and Future. *Information Sciences*, 179(8):1102–1115.

Reidenberg, Joel R., Jaspreet Bhatia, Travis Breaux, and Thomas Norton. 2016. Ambiguity in Privacy Policies and the Impact of Regulation. *Journal of Legal Studies*, 45:S163–S190.

Smith, H. Jeff, Tamara Dinev, and Heng Xu. 2011. Information Privacy Research: An Interdisciplinary Review. *MIS Quarterly*, 35(4):989–1016.

Schwartz, Paul M., and Daniel Solove. 2009. *Notice and Choice: Implications for Digital Marketing to youth. Memo Prepared for the Second NPLAN/BMSG Meeting on Digital Media and Marketing to Children*. Berkeley, CA: Berkeley Media Studies Group.

Schwartz, Alan, and Louis L. Wilde. 1979. Intervening in Markets on the Basis of Imperfect
Information: A Legal and Economic Analysis. *Faculty Scholarship Series*, 1117:630–682.

Strahilevitz, Lior Jacob, and Matthew B. Kugler. 2017. Is Privacy Policy Language Irrelevant to Consumers? *Journal of Legal Studies*, 45(52):S69–S95.

Tsai, Janice Y., Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. 2011. The Effect of Online Privacy Information On Purchasing Behavior: An Experimental Study. *Information Systems Research*, 22(2):254–268.

Turow, Joseph, Michael Hennessy, and Nora Draper. 2015. *The Tradeoff Fallacy: How Marketers are Misrepresenting American Consumers and Opening Them up to Exploitation*. Philadelphia, PA: Annenberg School for Communication, University of Pennsylvania. Available from https://www.asc.upenn.edu/sites/default/files/TradeoffFallacy_1.pdf

Varian., Hal R. 1996. Economic Aspects of Personal Privacy. in *U.S. Department of Commerce, Privacy and Self-Regulation in the Information Age*. Washington, DC: U.S. Department of Commerce, National Telecommunications and Information Administration.

Vila, Tony, Rachel Greenstadt, and David Molnar. 2003. Why We Can't be Bothered to Read Privacy Policies: Models of Privacy Economics as a Lemons Market. *ACM International Conference Proceeding Series, 5*:403–407.

Zimmeck, Sebastian, and Steven M. Bellovin. 2014. *Privee: An Architecture for Automatically Analyzing Web Privacy Policies*. San Diego, CA: 23rd USENIX Security Symposium.

Table 1

"What are the reasons why you usually do not read or read only partially the privacy statements?" (Multiple answers possible)

Variable	Description	Mean	Min	Max
NOT HONORED	Takes value 1 if individual finds the websites will not honor privacy policies in any case and 0 otherwise	0.148	0	1
PROTECT LAW	Takes value 1 if individual thinks the law protect him and 0 otherwise	0.144	0	1
FINDABILITY	Takes value 1 if individual does not know where to find the privacy policy and 0 otherwise	0.075	0	1
NOT IMPORTANT	Takes value 1 if individual believes it is not important to read privacy policies and 0 otherwise	0.116	0	1
TOO LONG	Takes value 1 if individual finds privacy policies are too long to read and 0 otherwise	0.627	0	1
UNCLEAR	Takes value 1 if individual finds privacy policies are unclear and 0 otherwise	0.355	0	1
POLICY PRESENT	Takes value 1 if individual finds the presence of privacy policy is sufficient and 0 otherwise	0.166	0	1
NUMBER OF OBSERVATIONS		13783		

Table 2

Clusters composition

Cluster	Variables	Description
1	TOO LONG	Understanding privacy policies
	UNCLEAR	
2	POLICY PRESENT	Trust of the institutions
	PROTECT LAW	
3	NOT IMPORTANT	Lack of interest in privacy policies (low privacy preferences)
	FINDABILITY	
4	NOT HONORED	Lack of trust of websites

Table 3

Descriptive statistics

Source: Own elaboration. The number of observations is 13783.

Variable	*Description*	*Mean*	*Std. Dev.*	*Min*	*Max*
KN_AUTHORITY	Takes value 1 if the individual knows the existence of right authorities about privacy, 0 otherwise	0.430	0.495	0	1
TRUSTGOVERNEMENT	Takes value 1 if the individual declares to trust on national government	0.452	0.497	0	1
MISUSE	Measures the intensity of privacy concern 1 indicates low level of privacy concern….and 4 indicates high level of privacy concern.	2.863	0.810	1	4
Socio-demographic					
AGE	Age of the individual	44.57	16.35	15	98
MALE	Takes value 1 if the individual is male	0.485	0.500	0	1
MARRIED	Takes value 1 if the individual is married, 0 otherwise	0.523	0.499	0	1
STUDENT	Takes value 1 if the individual is a student, 0 otherwise	0.095	0.293	0	1
DEGREE	Takes value 1 if the individual has a degree, 0 otherwise	0.422	0.494	0	1
SELF_EMP	Takes value 1 if the individual is a self-employed, 0 otherwise	0.085	0.279	0	1
EMPLOYED	Takes value 1 if the individual is an employee, 0 otherwise	0.501	0.500	0	1
MIDDLE	Takes value 1 if the individual lives in a middle town, 0 otherwise	0.423	0.494	0	1
LARGE	Takes value 1 if the individual lives in a large town, 0 otherwise	0.292	0.455	0	1
NODIFFICULTY	Takes value 1 if the individual declares to have no difficulty paying the bills by the end of the month	0.647	0.478	0	1
Use of internet					
FREQ_INT	Internet use frequency: 1 if never (..) 6 if daily	4.714	0.758	1	6

Table 4

Results of the cluster analysis with the group average linkage method

	Cluster 1 TOO LONG UNCLEAR	**Cluster 2** POLICYPRESENT PROTECT_LAW	**Cluster 3** FINDABILITY NOT IMPORTANT	**Cluster 4** NOTHONORED
Descriptive statistics of each cluster				
	Mean	Mean	Mean	Mean
KN_AUTHORITY	.451	.494	.395	.490
MISUSE	2.91	2.7444	2.74	3.03
TRUSTGOV	.435	.491	.409	.342
AGE	44.04	44.35	43.77	43.79
MALE	.481	.495	.496	.524
MARRIED	.509	.531	.531	.516
STUDENT	.099	.099	.099	.090
DEGREE	.434	.444	.346	.420
SELF_EMP	.086	.086	.082	.092
EMPLOYED	.511	.514	.483	.509
MIDDLE	.422	.426	.414	.403
LARGE	.297	.288	.293	.292
NODIFFICULTY	.661	.665	.561	.604
FREQINT	4.77	4.75	4.59	4.72

Table 5

Results of the logit estimations

	(1) Cluster 1 Too long, Unclear		(2) Cluster 2 Policy Present, Protect law		(3) Cluster 3 Findability, Not important		(4) Cluster 4 Not honored	
KN_AUTHORITY	0.020	(0.041)	0.223***	(0.040)	-0.199***	(0.048)	0.216***	(0.051)
MISUSE	0.261***	(0.026)	-0.209***	(0.025)	-0.206***	(0.028)	0.302***	(0.033)
TRUSTGOV	-0.049	(0.043)	0.291***	(0.042)	0.042	(0.049)	-0.360***	(0.054)
AGE	-0.012***	(0.002)	-0.001	(0.002)	0.001	(0.002)	-0.001	(0.002)
MALE	-0.116***	(0.040)	0.027	(0.039)	0.130***	(0.046)	0.230***	(0.049)
MARRIED	-0.060	(0.044)	0.040	(0.043)	0.021	(0.050)	-0.068	(0.053)
STUDENT	-0.016	(0.095)	0.082	(0.092)	-0.081	(0.104)	-0.066	(0.118)
DEGREE	0.057	(0.046)	0.048	(0.044)	-0.231***	(0.054)	0.118**	(0.056)
SELF_EMP	0.151*	(0.080)	0.042	(0.077)	-0.117	(0.090)	-0.035	(0.096)
EMPLOYED	0.125**	(0.049)	0.049	(0.049)	-0.080	(0.056)	-0.007	(0.062)
MIDDLE	0.003	(0.049)	0.001	(0.048)	0.023	(0.056)	-0.140**	(0.061)
LARGE	0.097*	(0.054)	-0.067	(0.053)	-0.038	(0.062)	-0.080	(0.066)
NODIFFICULTY	0.036	(0.046)	-0.004	(0.046)	-0.221***	(0.052)	-0.102*	(0.057)
FREQINT	0.218***	(0.026)	0.072**	(0.028)	-0.131***	(0.028)	0.053	(0.034)
_cons	0.310	(0.209)	-1.408***	(0.208)	-0.651***	(0.232)	-2.922***	(0.261)
Country fix effects	Yes		Yes		Yes		Yes	
N	13783		13783		13783		13783	

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01