Understanding Privacy Policies: Content, Self-Regulation, and Markets

Florencia Marotta-Wurgler[1]

NYU School of Law

January 3, 2016

The current regulatory approach to consumer information privacy is based on a "notice and choice" self-regulation model, but commentators disagree on its impact. I conduct a comprehensive empirical analysis of 261 privacy policies across seven markets and measure the extent to which they comply with the self-regulatory guidelines of the Federal Trade Commission (FTC), US-EU Safe Harbor Agreement, and others. I track terms involving notice, data collection, sharing, enforcement, security, and other practices, and create a measure of substantive protections. The average policy complies with 39% of the FTC guidelines issued in 2012, and there is no evidence that firms have updated their policies in response to these guidelines. Terms that require firms to bear costs or constrain their behavior are less likely to be included. Protections vary widely across markets, however: Adult sites offer the clearest notice of practices and report less data collection and sharing than other sites, while cloud computing firms report more extensively on data security practices. Overall, the results suggest that privacy policies are being shaped as much by market forces as by the current regulatory regime.

[1] New York University School of Law. I would like to thank Daniel Svirsky and Robert Taylor for outstanding work on the project, Oren Bar-Gill, Omri Ben-Shahar, Emiliano Catan, Kevin Davis, Chris Hoofnagle, William Hubbard, Louis Kaplow, Kirsten Martin, Helen Nissembaum, Katherine Strandburg, Ira Rubinstein, Lauren Willis, Jeff Wurgler, Kathryn Zeiler, participants at the Privacy Law Research Scholars Conference, Conference on Empirical Legal Studies, Boston University Law and Economics Workshop, Northwestern University Law and Economics workshop, University of Michigan Law and Economics workshop, the University of Chicago Law School conference on "Contracting Over Privacy," and Harvard Law School Faculty workshop, for helpful comments and suggestion on an earlier version of this draft. I would also like to thank Amanda Conley, Nicolas Heliotis, Alex Lipton, Julienne Markel, Isaac Sasson, Luke Smith, Melissa Quartner, Christopher Van Zele, and JingJing Wu for outstanding research assistance.

I. Introduction

Billions of people use the Internet every day to read the news, check email, connect with friends on social networks, buy groceries, to use a search engine to answer a particular question or find a site or document, and so on. Every keystroke and mouse-click flows into a stream of information on that individual's characteristics, needs, wants, finances, address, family, friends, and more. Companies often collect this information for commercial purposes, including constructing user-specific profiles to target content or advertising, enhance the services they offer, or to share or sell it on to third parties to do the same.[2] In most cases consumers have little ability to control or even keep track of this information without abandoning the Internet altogether.

It is both obvious and documented that consumers care about their information privacy in certain contexts, and they are concerned about misuses or leaks of such information.[3] Policymakers are continually considering revisions to the current regulatory model of information protection.[4] For the most part, consumer information has been protected by a self-regulatory regime articulated by the Federal Trade Commission (FTC). This model, known as "notice and choice," has been predominantly based on disclosure and has encouraged firms to adopt substantive information protections via self-regulation.[5] It asks that companies adopt privacy policies which disclose their practices related to the collection, use, sharing, and security of consumer information, and that these practices conform to a set of "Fair Information Practices" (FIPs).[6]

The goal of this regulatory model has been to facilitate "competition on privacy" by allowing consumers to inform themselves, either by reading policies or relying on third party certification seals, and choose which sites to use based on this knowledge. In theory, disclosure can alleviate market failures that stem from asymmetric information while preserving consumer choice. Self-regulation is also desirable, in theory, as it allows firms flexibility and the ability to determine cost effective ways of compliance.

Self-regulation works only if industry participants actually develop and adopt appropriate codes of conduct following the FTC guidelines. But the FTC has limited ability to encourage this: it has little rule-making authority and cannot impose sanctions

---

[2] Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL L. REV. 1503.

[3] *See* Section II.

[4] *See, e.g.*, Data Broker Transparency and Accountability Act, S. 2025, 113th Cong. (2014); Data Security and Breach Notification Act of 2014, S. 1976, 113th Cong.; Personal Data Privacy and Security Act of 2014, H.R. 3990, 113th Cong.; Geolocational Privacy and Surveillance Act, H.R. 1312, 113th Cong. (2013); Email Privacy Act, H.R. 1852, 113th Cong. (2013); Location Privacy Protection Act of 2014, S. 2171, 113th Cong.; Eliminate Privacy Notice Confusion Act, H.R. 749, 113th Cong. (2013); Alexis Agin Identity Theft Protection Act of 2013, H.R. 2720, 113th Cong.

[5] *See infra* Section II and accompanying text. A number of state laws also protect information privacy in certain contexts. *See, e.g.*, CALIF. BUS. & PROF. CODE §§ 22575–22578 (requiring website operators to post privacy policies describing their information practices); CONN. GEN. STAT. § 42-471 (requiring businesses who collect Social Security information in the course of their business to implement privacy protection policies); NEBRASKA STAT. § 87-302(14) (prohibiting firms from making false or misleading statements in privacy polices).

[6] FTC, *Privacy Online: A Report to Congress* 7 (1998), http://www.ftc.gov/sites/default/files/documents/public_events/exploring-privacy-roundtable-series/priv-23a_0.pdf. California law also requires that firms adopt privacy policies giving notice of their privacy practices. *See* Cal. Bus. & Prof. Code §§ 22575–22577.

to actors for noncompliance. This approach has consequently been labeled as incomplete and toothless, and incapable of alleviating potential market failures.[7]

On the other hand, some prominent commentators have recently argued that the existing approach is more substantive and effective than previously thought. Their view is that the FTC has effectively encouraged self-regulation and the adoption of reasonable information practices by firms by threatening to push for regulation for failure to comply; by giving the regime some teeth by policing and enforcing any violations of privacy policies and codes of conduct under Section 5 of the FTC Act, which targets unfair and deceptive practices; by making its privacy actions and initiatives public; and, by encouraging the development of certification regimes.[8] Further, current regulatory approaches continue to embrace self-regulation (to work in tandem with the adoption baseline privacy protections, which to date have not been adopted). Several bills seeking to protect consumer information privacy continue to be drafted or considered.

Unfortunately, regulatory discussions have taken place in the context of media anecdotes about practices of a handful of large firms and a few scattered, small-sample studies, not systematic evidence about the actual content of typical privacy policies or the adoption or effectiveness of notice and choice models.[9] This paper provides the first large-sample study of the content of modern privacy policies, their degree of compliance with self-regulatory guidelines, and the protections they offer. More generally, it offers empirical insight into a model of consumer protection that seeks to go beyond regulation, a popular regulatory tool that has mostly failed to achieve its objectives.

I analyze 261 privacy policies across seven online markets where people share personal and personally identifiable information to different degrees—adult sites, cloud computing, dating sites, gaming sites, news and reviews, social networks, and forum/special interest sites. These are services where consumers often provide private information, sometimes highly sensitive information, and have reason to be concerned with privacy practices, so they make for an interesting and rich sample for study.

For each policy, a group of research assistants and I hand-coded the presence or absence of 49 different terms or practices. These terms address many aspects of privacy practices, from giving notice of the types and uses of data the firm collects to the internal security practices used to protect that information. These 49 terms were chosen because they appeared in at least one guideline that has been influential or that govern current consumer information practices, including those introduced in 2012 by the FTC, *Protecting Consumer Privacy in an Era of Rapid Change*, the FTC's prior guidelines from 2000, the principles suggested by the White House in 2012 in its *Consumer Privacy Bill of Rights*, and the original privacy information privacy guidelines of a report by the Department of Housing, Education, and Welfare Fair Information Practices of 1973 (1973 FIPS), which influenced all subsequent guidelines. I also track compliance with the United States-European Union Safe Harbor Agreement (US-EU Safe Harbor), a voluntary framework which until quite recently allowed US firms interacting with EU

[7] *See infra* Section II.
[8] Id.
[9] An exception to this is early surveys on information privacy practices conducted around 2000 and cited by the FTC in its report to Congress. *See infra* Section II.

citizens a safe harbor in complying with the EU data privacy laws.[10] Measuring policies against these benchmarks allow me to evaluate the degree to which firms adhere to them, as required.

The results give a sobering picture of compliance with guidelines. The average policy in the sample complies with only 39% the 2012 FTC guidelines, the operational regime for the firms in our sample, although drafts have been circulating since 2010. Only 66 out of 261 policies comply with more than half of the 2012 FTC guidelines. Also, only 12 out of 56 policies *that claim to adhere to US-EU Safe Harbor requirements* actually contain text that complies with at least half of its requirements.

Per the notice and choice paradigm, several guidelines involve giving notice of practices. But despite an average length of 2,176 words—most firms do not follow FTC recommendations to adopt short, streamlined, and standardized privacy policies—policies are often silent on crucial or required categories. Silence is problematic in this context because there are no clear gap-filling default rules akin to those in the UCC's Article 2. Judicial opinions, some information statutes, and FTC enforcement actions provide some guidance, but there are a number of categories of terms where the meaning of silence is unclear.[11] An additional complication is that when terms are included in the policy, they are not infrequently contradictory or ill-defined.

In terms of substance, data collection practices are extensive and often appear to violate regulatory guidelines. The overwhelming majority of policies report that they collect contact information, computer information (such as IP address and browser type) and interactive information (such as browsing behavior or search history). Relatively few firms claim to limit the use of personally identifiable information to internal or context-specific purposes. More than two-thirds state that they share information with third parties, but do not report having a contract with those parties to limit the use of the shared data or bind those parties to its own privacy policy. Simply put, many companies in the sample collect a lot of information but consumers have no way to know where it goes, how it is used, or whether the chain of custody even ends.

An interesting question is whether "compliance," in the sense of notice of practices and substantive protections, such as security and respect for context, would be even more modest in the absence of guidelines. This is an important question because even if compliance is imperfect, the current regulatory model would clearly be better than no guidelines at all. In other words, I am interested in the extent to which the terms firms offer can be interpreted as deliberate efforts to satisfy guidelines or would be offered anyway as a result of consumer demand and market forces. It is a difficult question to answer, but several patterns in the data suggest that at least some of the compliance with guidelines could best be described as incidental.

In particular, privacy protections and compliance differ across markets in intuitive and robust ways. In terms of overall compliance with FTC guidelines, adult sites stand

---

[10] The US-EU Safe Harbor Agreement was invalidated in October, 2015 by the European Court of Justice in Case C-362/14, Schrems v. Data Prot. Comm'r, 2015 E.C.R. ___, *available at* http://curia.europa.eu/juris/celex.jsf?celex=62014CJ0362&lang1=en&type=TXT&ancre= (finding that U.S. surveillance practices violate the privacy rights of EU citizens).

[11] 15 U.S.C. § 45. For a full discussion of the FTC cases under Section 5 of the FTC Act, see Daniel Solove & Woodrow Hartzog , *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583 (2014).

out, especially on some of the most important dimensions. Adult sites are more likely to comply with FTC notice requirements; their policies often contain clear and concise descriptions of the collection and uses of data. Adult sites are also likelier to comply with the FTC's data sharing guidelines, which impose constraints on how and when the information should be shared. Basically, adult sites effectively communicate to users that they collect and share relatively little data, which is presumably what users prefer. If there is one market where individuals might be aware and concerned about their information privacy, it might be this one.

Another robust market difference involves cloud computing sites, which are far more likely than firms in other markets to comply with data security requirements. They are also more likely to implement substantive information protection practices throughout their organization, so-called privacy by design practices. This would also appear to be a natural outcome of market competition. Cloud computing users entrust numerous and important files to the firm only because they expect them to be preserved and protected. Cloud computing firms thus have a special interest in advertising their security measures.

Finally, I compare policies that were updated after 2012, when the new FTC guidelines were finalized, with policies last updated before 2012. If firms are deliberately responding to FTC guidelines, one would expect to see compliance with more of the new guidelines than the old ones. This is not the case. The fresher policies do not appear to be written with any special eye to the current guidelines.

The results offer a more nuanced picture of the factors that might be shaping the content of privacy policies. They also suggest that the current regulatory environment could be fertile ground for consumer abuse. The idea behind "competing on privacy" is that consumers can shop around for websites whose privacy policies are satisfying. But when policies are complex, inconsistent, and incomplete, and subject to only minimal third party certification, the consumer has no way of knowing what data are collected and where they go. Further, there are unclear default rules in the area. More realistically, given that consumers tend to ignore fine print, a more serious implication of this complexity is that it is difficult for intermediaries to simplify the terms in current policies in a consumer-friendly way or convert them into machine-readable policies able to be standardized or personalized. Given the known failures of disclosure regulation, a more important finding is that firms' relatively weak embrace of self-regulation may have resulted in scant adoption of substantive information privacy protections. It is true that when privacy concerns are especially salient, such as the adult market, competitive market forces appear to be shaping policies in ways consistent with consumer preferences. But for the broader mass of policies these forces may be weak. More generally, the analysis could help inform current debates regarding consumer information privacy and the desirability of self-regulatory regimes.

One limitation of the analysis is that I cannot examine company practices beyond those disclosed in privacy policies. Policies might suggest unseemly practices but behavior might nonetheless be constrained by alternative disciplining mechanisms, such as reputation or enforcement actions under state laws or the FTC's "Unfair and Deceptive Practices" Act.[12] Or, policies might include protective terms that have no teeth in practice.

[12] *But see* Florencia Marotta-Wurgler & Daniel Svirsky, *infra* note 59 (evaluating the deterrent effect of FTC privacy enforcement actions under Section 5 by examining changes in firms' privacy policies).

All I can measure is how much the text of the privacy policy complies with the text of regulatory guidelines. But regulators are in the same position. They, too, must base their activities on what they can measure. Disclosure, in particular, has been an important component of notice and choice regulation for decades and FTC enforcement actions have often centered on the statements made within policies.

The paper proceeds as follows. Section II offers background on the laws governing information privacy online and reviews the prior literature. Section III explains the sample and presents the main analysis of contract content and degrees of compliance with guidelines. Section IV investigates the extent to which observed compliance can be attributed to deliberate efforts to satisfy the guidelines versus incidental compliance that firms might have offered anyway. Section V discusses some implications and concludes.

II. Background

Individuals today share extensive amounts of information online with apparent comfort. People post updates on social networks, display their employment credentials on LinkedIn for networking purposes, and are happy to receive movie recommendations that Netflix produces by analyzing past movie choices. Individuals have also embraced the convenience of online commerce and the highly dynamic world of mobile applications that offer a wide variety of useful services and information. Most of these entities collect, use, and share their personal information with others, including third parties unrelated to the original transaction who then use this information for secondary purposes.

In addition to the normative considerations regarding such data collection and personalization,[13] a major question in debates about the regulation of information privacy has been whether individuals care about information privacy or hold beliefs about that are inconsistent with current practices. Survey and experimental evidence offer some support for both. A Pew Research Center survey of 2013 reports that 50% of respondents are concerned about the amount of information about them available online and that 68% believe that current laws do not adequately protect information privacy.[14] Other surveys find similar results and document an increased reluctance by consumers to share private information over time.[15]

---

[13] See, e.g., Yochai Benkler, Net Regulation: Taking Stock and Looking Forward, 71 U. COLO. L. REV. 1203, 1253 (2000) ("From the normative perspective, such a development undermines individual autonomy because it pervasively displaces personal control over the information environment within which individuals view the world, because the perception of the world and of possible options for action are defined by others."); see also Alessandro Acquisti, Curtis R. Taylor & Liad Wagman, The Economics of Privacy (Working Paper, 2015), available at http://ssrn.com/abstract=2580411 (explaining the tradeoffs associated with the protection and disclosure of personal information).

[14] See, e.g., PEW Survey, Privacy, Anonymity, and Security Online (September 5, 2013), available at http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online.

[15] Avi Goldfarb & Catherine Tucker, Shifts in Privacy Concerns, 102 AM. ECON. REV. 349, 353 (2012); Alessandro Acquisti & Jens Grossklags, Privacy and Rationality in Individual Decision Making, 3 SECURITY & PRIVACY 26 (2005); see also Chris Hoofnagle, Jennifer King, Su Li & Joseph Turow, How Different Are Young Adults from Older Adults When it Comes to Information Privacy Attitudes and Policies? (Apr. 14, 2010), http://ssrn.com/abstract=1589864 (reporting representative telephone survey

Laboratory experiments have found that consumers value privacy in context-dependent ways and that consumers would sometimes be willing to pay to protect their information from being disclosed.[16] Other studies show that consumers have incorrect beliefs about how much of their information is collected and shared and have difficulty understanding the meaning of privacy policies. All of this research shows that the tradeoffs involved in sharing personal information are often complicated, as they are often context-specific and bundled with other products or services, which in turn makes it difficult to make welfare-maximizing choices about privacy.[17]

A. Notice and Choice Self-Regulation

Discussions regarding the regulation of information collection practices by commercial entities began in earnest in the 1990s as such practices became pervasive.[18] At that time, e-commerce was nascent and Congress was reluctant to risk derailing or hamper its development. Instead, Congress charged the FTC to create guidelines of fair information practices for firms to adopt through self-regulation and encourage the development of market mechanisms to increase consumer information protection.[19]

---

results of 1000 individuals in the United States showing that younger adults have similar privacy concerns and desires of information protection as older adults); Joseph Turow, Chris Jay Hoofnagle, Deirdre K. Mulligan, Nathaniel Good & Jens Grossklags, The Federal Trade Commission and Consumer Privacy in the Coming Decade, 3 I/S J.L. & POL'Y INFO. SOC'Y 723, 729–32 (2007) (collecting survey evidence revealing that consumers worry about online privacy and believe privacy policies are meant to protect them).

[16] Alessandro Acquisti, Leslie K. John & George Loewenstein, What is Privacy Worth?, 42 J. LEGAL STUD. 249 (2013); Janice Y. Tsai, Serge Egelman, Lorrie Cranor & Alessandro Acquisti, The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study, 22 INFO. SYS. RES. 254 (2007). Lior Strahilevitz and Matthew Kugler, Is Privacy Policy Language Irrelevant to Consumers? (finding that a non-trivial fraction of respondents would be willing to pay to protect their privacy) (working paper, 2015).
See also Alex Marthews & Catherine Tucker, Government Surveillance and Internet Search Behavior (Apr. 29, 2015), http://ssrn.com/abstract=2412564 (finding a statistically significant impact on users' Google search behavior after the Snowden revelation, albeit of a magnitude of around two percent).

[17] Acquisti & Grossklags, supra note 15

[18] See Benkler, supra note 13, at 1247 (explaining how in the 1990s Congress focused on regulation to respond to structural changes in information flows); see also JONATHAN ZITTRAIN, THE FUTURE OF THE INTERNET AND HOW TO STOP IT (2008).

[19] The protection of information privacy in the United States has followed an area-specific approach, where specific statutes govern particular sectors of information privacy. See, e.g., Children's Online Privacy Protection Act of 1998 (COPPA), Pub. L. No. 105-277, § 1302, 112 Stat. 2681-728 (codified at 15 U.S.C. §§ 6501–6506) (governing the collection of information of children 13 and under), Gramm-Leach-Bliley Ac (GLBA) 15 U.S.C. §§ 6801–6809 (2012) (governing the collection and use of financial information). While Congress enacted laws regulating information privacy in particular sectors, states, in particular California, actively pursued innovative regulations on the information privacy front. For example, California's Data Breach Notification Law of 2002 became a model for other states, most of which followed suit. CAL. CIV. CODE §§ 1798.29, 1798.82 (West 2008). California also enacted the California Privacy Protection Act, which entitles consumers to find out how their personal information is shared by companies for marketing purposes and encourages companies to allow consumers to opt-out of such sharing. CAL. CIV. CODE § 1798.83 (West 2008). Other states, such as Nebraska and Pennsylvania, enacted laws prohibiting companies from making false or misleading statements in their privacy policies. NEB. REV. STAT. § 87-302(1) (2015); 18 PA. STAT. AND CONS. STAT. ANN. §4107(a)(1).

The FTC developed its initial set of guidelines in a report to Congress in 1998. These were based on the principles of an influential 1973 report by the Department of Health, Education, and Welfare on the protection of information privacy, *Records, Computers, and the Rights of Citizen* [1973 HEW FIPs] outlining practices including limited collection and use of information, access to the individual, security safeguards, and accountability.[20] In its report, the FTC encouraged firms to adopt privacy policies describing their data practices, including collection, use, sharing, and security of personal information collected, and to give choices to consumers regarding certain collection and uses of information. This model became known as "notice and choice." Howard Beales, former Director of the FTC Bureau of Consumer Protection, explained it as follows: "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."[21]

Over the next fourteen years, the FTC took a leading role by continuing to develop information privacy guidelines and encouraging their adoption by bringing actions against firms who violated their privacy commitments or engaged in unfair practices under Section 5 of the FTC Act, which gives the FTC the power to police "unfair and deceptive" trade practices.[22] The guidelines consist of those outlined in a 2000 report to Congress, *Protecting Consumer Privacy in an Era of Rapid Change* [FTC 2000], and a revised set outlined in a 2012 report to Congress, *Protecting Consumer Privacy in an Era of Rapid Change* [FTC 2012]. The FTC was also charged with enforcing the US-EU Safe Harbor Agreement (US-EU Safe Harbor), another voluntary framework enacted in the year 2000 that allows firms interacting with EU member states' citizens to comply with the EU's more stringent data requirements.[23] The SHA allows US firms who voluntarily adhere to its requirements to qualify as offering adequate protections for personal information collected from EU citizens. Firms seeking to abide by the SHA must have privacy policies satisfying its requirements.

Like most disclosure regimes,[24] a potential limitation with notice and choice is that consumers would just ignore the privacy policies, or that the policies would be

---

[20] See U.S. Dep't of Health, Educ. & Welfare, Records, Computers, and the Rights of Citizens (1973), available at http://www.justice.gov/opcl/docs/rec-com-rights.pdf.

[21] Howard Beales, Dir., Bureau of Consumer Prot., Remarks on the Privacy Notices and the Fed. Trade Comm'n's 2002 Privacy Agenda (Jan. 24, 2002), available at https://www.ftc.gov/public-statements/2002/01/privacy-notices-and-federal-trade-commissions-2002-privacy-agenda.

[22] 15 U.S.C. §45. During the past decade and a half, the FTC has brought about more than 100 actions against firms under deception for breaching terms in their privacy policies and for engaging in unfair practices. A summary of the actions can be found at https://www.ftc.gov/news-events/media-resources/protecting-consumer-privacy/enforcing-privacy-promises.

[23] Council Directive 95/46, on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, art. 2(a), 1995 O.J. (L 281) (The SHA was entered into to ease compliance by US firms with foreign information privacy laws, such as Article 25 of the European Union Data Directive, which limits transfer of personal data to countries lacking "adequate" levels of protection

[24] See Omri Ben-Shahar & Carl Schneider, supra note 5 (providing a thorough review of the failures of mandated disclosure); Florencia Marotta-Wurgler, Will Increased Disclosure Help? Evaluating the Recommendations of the ALI's "Principles of the Law of Software Contracts," 78 U. CHI. L. REV. 165 (2011). But see Oren Bar-Gill, SEDUCTION BY CONTRACT: LAW, ECONOMICS AND PSYCHOLOGY IN

inaccessible.[25] Indeed, in a 2010 preliminary report to Congress, the FTC found just that and stated that "[t]he Notice and Choice model, as implemented, has led to long, incomprehensible privacy policies that consumers typically do not read, let alone understand."[26] The report did note that "[a]n important legacy of the Commission's notice-and-choice approach to privacy is that most companies now disclose their data practices to consumers through privacy notices. Indeed […] privacy notices continue to promote companies' accountability for their practices."[27] In its 2012 report, The FTC urged companies adopt smarter disclosures and increase transparency by making policies clearer, shorter, and standardized.

Attempting to address the readership issue early on, the FTC encouraged the development of third-party certification mechanisms, such as TRUSTe and BBB Online, that would enable adhering firms to display seals certifying the firms' adherence to some core privacy practices, as revealed in their privacy policies. In its report to Congress in 2012, the FTC further encouraged the role of third party information intermediaries by encouraging the standardization of policies to enable machine reading techniques for third parties to summarize and convey information in more accessible ways.

Despite the widespread adoption of privacy policies, many were not satisfied with notice and choice.[28] In addition to complex policies, the lack of rigorous enforcement mechanisms would likely render this approach toothless and likely to create collective action problems that would result in low compliance or in compliance just high enough to prevent a major overhaul.[29] Indeed, in its 2012 report, even the FTC expressed the need

---

CONSUMER MARKETS (Oxford University Press, 2012) (presenting more sophisticated forms of disclosure regulation aimed at simplifying the presentation of information to consumers); Cass R. Sunstein & Richard H. Thaler, Libertarian Paternalism is Not an Oxymoron, 70 U. CHI. L. REV. 1159 (2003) (explaining how certain kinds of disclosures can improve decision making).

[25] See Yannis Bakos, Florencia Marotta-Wurgler, and David R. Trossen, Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts, 43 J. LEGAL STUD. 1 (2014) (reporting that only 0.1% of consumers read software End User License Agreements).

[26] FTC, Protecting Consumer Privacy in an Era of Rapid Change: Recommendations For Businesses and Policymakers, [hereinafter, FTC 2010 Preliminary Report] 2010, available at https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-bureau-consumer-protection-preliminary-ftc-staff-report-protecting-consumer/101201privacyreport.pdf; see also Alicia McDonald & Lorrie Cranor, The Cost of Reading Privacy Policies, 4 I/S J.L. & POL'Y INFO. SOC'Y 540 (2008) (estimating that it would take an average 244 hours per year for each individual to read the privacy policies of each web site visited once a month for a total cost of $3,534 a year)

[27] See FTC 2010 Preliminary Report, supra note 26, at 70 (citing remarks by Fred Cate, Paula Breuning, and a written comment of the Business Forum for Consumer Privacy). The report also stated that "The public posting of privacy notices is especially valuable to consumer privacy advocacy groups, regulators, and those consumers who want to learn more about a company's overall privacy practices." See Richard Craswell, Static Versus Dynamic Disclosures, and How Not to Judge Their Success or Failure, 88 WASH. L. REV. 333 (2013) (explaining the different functions of static and dynamic disclosures).

[28] See e.g., Schwartz, supra note 29; Kang, supra note 29; Fred Cate, The Limits of Notice and Choice, 8 IEEE SEC. & PRIVACY 59, 59–62 (2010); Kirsten E. Martin, Transaction Costs, Privacy, and Trust: The Laudable Goals and Ultimate Failure of Notice and Choice to Respect Online Privacy, FIRST MONDAY (Dec. 2013). But see M. Ryan Calo, Against Notice Skepticism in Privacy (and Elsewhere), 87 NOTRE DAME L. REV. 1027 (2013).

[29] See, e.g., Benkler, supra note 13 ("Almost certainly, however, in the absence of regulation, the digitally networked environment will be significantly more subject to surveillance than the analog environment - because it can be, and because the constraints will only be placed to reach a level just below the threshold

for Congress to adopt baseline privacy regulations to work in tandem with firms' self-regulatory efforts.[30] But the report outlined a new and revised set of guidelines, encouraging firms to adopt more substantive information practices to be incorporated in all aspects of management and product development, such as employee training and minimizing collection and use of personal information. It labeled this self-regulatory approach "privacy by design."

Yet some have embraced the current approach, and some commentators, including the FTC in its earliest report to Congress,[31] have recently favored it as a less disruptive mechanism to encourage firms to protect information privacy in a cost-effective and flexible manner.[32] In their view, a number of factors contribute to the relative success of this approach, including the FTC's threat to encourage Congress to regulate if self-regulation is not embraced, as well as its now extensive body of privacy enforcement actions, whose consent decrees offer additional input into which information practices the FTC considers adequate and which it does not. An example cited for this is the widespread adoption of privacy policies.[33]

The FTC guidelines, SHA requirements, and other state statutes (such as California's Online Privacy Protection Act) presumably also help to account for the existence of privacy policies.[34] The FTC guidelines can only encourage their adoption, not require it. Indeed, given that most FTC enforcement actions under Section 5 have

---

of consumer rebellion, but no lower."); see also Jerry Kang, Information Privacy in Cyberspace Transactions, 50 STAN. L. REV. 1193, 1253 (1998); Paul M. Schwartz, Internet Privacy and the State, 32 CONN. L. REV. 815, 833 (2000) (arguing that the proposed regulatory scheme provides weak incentives for firms to comply because there are no clear enforcement mechanisms); Lior Strahilevitz, Toward a Positive Theory of Privacy Law, 126 Harvard L. Rev 2010 (2013) (discussing the distributive dimensions of privacy regulation).

[30] See FTC, Protecting Consumer Privacy in an Era of Rapid Change ii (2012) ("The commission agrees that, to date, self-regulation has not gone far enough…. [E]ven in some well-established markets, basic privacy concepts like transparency about the nature of companies' data practices and meaningful consumer choice are absent. This absence erodes consumer trust.") at 12.

[31] See FTC, Self-Regulation and Privacy Online: Report to Congress 6 (July, 1999), available at https://www.ftc.gov/system/files/documents/reports/self-regulation-privacy-onlinea-federal-trade-commission-report-congress/1999self-regulationreport.pdf. (Referring to self-regulation as "the least intrusive and most efficient means to ensure fair information practices, given the rapidly evolving nature of the Internet and technology.")

[32] See Ira Rubinstein, Privacy and Regulatory Innovation: Moving Beyond Voluntary Codes 6 I/S: J.L. & POL'Y INFO. SOC'Y 356 (2011) (offering a comprehensive review of the arguments regarding the desirability of self-regulation); Peter P. Swire, Markets, Self-Regulation and Government Enforcement in the Protection of Personal Information, in PRIVACY AND SELF-REGULATION IN THE INFORMATION AGE 3 (U.S. Dep't of Commerce ed., 1997) (arguing that that the threat of regulation should provide a sufficient incentive for firms to comply with self-regulatory standards).

[33] Kenneth A. Bamberger & Deirdre K. Mulligan, Privacy on the Books and on the Ground, 63 STAN. L. REV. 247 (2010) at 288 ("Furthermore, [FTC] persuasion was critical in encouraging companies operating online to post privacy policies.") But see *See* A. Mitchell Polinsky & Steven Shavell, The Economic Theory of Public Enforcement of Law, 38 J. ECON. LIT. 45, 78 (2000) (noting that settlements reduce deterrence by stunting the development of legal precedent, which might in turn reduce deterrence).

[34] The Online Privacy Protection Act of 2003, Cal. Bus. & Prof. Code §§ 22575-22579 (2004) (The requirements of this law are not as exhaustive as the FTC privacy guidelines, so they cannot fully explain all terms that we see. The act requires web site operators that collect personally identifiable information to place conspicuous privacy policies, and offer some disclosure, including the type of information collected and the type of third parties with whom it shares information, among others)

focused on misleading statements made in privacy policies, one might expect that firms would choose either to not have privacy policies unless required by law or, if they did, to write into them few or no commitments.[35] Yet these commentators consider Section 5 and SHA enforcement actions as central to the success of the current approach.[36] This, together with other active privacy roles the FTC has taken, has led those who embrace the current regime to conclude that "[t]he FTC has essentially turned a mostly self-regulatory regime into one with some oversight and enforcement."[37]

Current approaches continue to rely on self-regulation. The 2012 FTC guidelines encourage firms to embrace privacy by design and other substantive protections (in conjunction with baseline regulation, which has not yet been adopted). In addition, the White House 2012 report on consumer information privacy, *A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy*, calls for multi-stake holder regulation.[38] The report embraced and expanded the approach of the 2012 FTC guidelines and created a Consumer Privacy Bill of Rights [WH Privacy Bill of Rights]. While a number of privacy bills were introduced in Congress in the wake of the report, none has yet made it into law.[39] A number of new bills wait in Congress and state legislatures.[40] For now, the protection of consumer information privacy remains mostly within the self-regulatory regime, yet empirical evaluations of this approach are mostly lacking.

.

B. Five Sets of Guidelines and a Theoretical Benchmark

---

[35] See Bamberger and Mulligan supra note 33 at 288 ("T]he publication of company policies making representations about practices with respect to personal information became central to the [FTC]'s initial exercise of its Section 5 enforcement jurisdiction, because the least controversial manner for the FTC to exercise its authority in the privacy area was to address factually misleading claims.") See also Solove and Hartzog supra note 11 at 594 ("Privacy policies were largely a voluntary measure by companies on the Internet to promote their privacy practices and partially an attempt at self-regulation in order to stave off further regulation…To a significant extent, the approach was successful.")

[36] See Solove and Hartzog supra note 11at 604 ("FTC enforcement serves as the lynchpin to the Safe Harbor Agreement, and Section 5 privacy enforcement serves as the lynchpin that makes the U.S. self-regulatory approach more than hollow…[T]he FTC has filled a great void, and without the FTC the U.S approach to privacy regulation would lose nearly all of its legitimacy.")

[37] Id at 2064. See also Steven Hetcher, The FTC as Internet Privacy Norm Entrepreneur, 53 Vand. L. Rev. 2041 (2000) at 2045 (arguing that firms pay attention to FTC reports); Bamberger and Mulligan at 287 ("Central to the FTC's emerging role as privacy regulators was its employment of regulatory tools outside the enforcement context, notably publicity, research, best-practice guidance, the encouragement of certification regimes, the enlist of expert input, and other deliberative and participatory processes promoting dialogue with advocates and industry.")

[38] The White House, Consumer Data Privacy in a Networked World 47 (2012), available at https://www.whitehouse.gov/sites/default/files/privacy-final.pdf.

[39] See supra note 4 and accompanying text; see also The Commercial Privacy Bill of Rights of 2011, S. 799, 112th Cong.; The Location Privacy Protection Act of 2011, S. 1223, 112th Cong. But see The Cybersecurity Information Sharing Act, S. 754, 114th Cong. (2015), which passed in the Senate in 2015.

[40] See Consumer Privacy Protection Act, S. 1158, 114th Cong. (2015); Student Digital Privacy and Parental Rights Act of 2015, H.R. 2092, 114th Cong.; Data Broker Accountability and Transparency Act, S. 668, 114th Cong. (2015).

I focus on the terms outlined or implied in five influential and relevant sets of guidelines, 1973 HEW FIPs, FTC 2000, US-EU Safe Harbor, FTC 2012, and WH Privacy Bill of Rights. Why these guidelines? The 1973 HEW FIPs provided the foundation of subsequent guidelines and thus provide a natural theoretical benchmark, even if subsequent guidelines don't fully incorporate all its principles. The FTC 2000 Report to Congress outlined the first set of regulatory guidelines that informed consumer information privacy practices until 2012, when the FTC presented a set of renewed guidelines to Congress. Both sets of guidelines may have shaped the information privacy practices of firms since the explosion of Internet commerce. The same is the case for the US-EU Safe Harbor Agreement. In addition, these guidelines, as well as the US-EU Safe Harbor Agreement, have been the root of many FTC Section 5 enforcement actions. I include the WH Privacy Bill of Rights to see the extent to which current policies adhere to the latest set of recommendations from a new source. Below is a brief overview of the principles behind each guideline.

B.1. HEW 1973 Fair Information Practice Principles

The report outlining fair information practices and recommending Congress to adopt them arose in response to the development of electronic databases and increased record-keeping systems. The report outlined five principles: there should be no secret data recording systems; individuals should be given means to find out what information is collected about them and how it is being used; individuals should be given means to prevent information collected for being used for a purpose other than the one originally collected without consent; individuals should be able to correct personal information about themselves; and, organizations creating, maintaining, or using personal information must assure the reliability of the data for its intended use and take precautions to avoid misuse of the data.

B.2. FTC 2000 Privacy Guidelines

The FTC 2000 focused on a subset of the 1973 FIPs and recommended that commercial websites collecting personal information comply with four basic principles: Notice, Choice, Access, and Security. To comply with the notice principle, firms are asked to provide consumers clear and conspicuous disclosures with the firms' information practices, including information related to collection, use, and sharing (including modes of collection, disclosures to third parties, and whether third parties can collect information entered on the firms' sites). To comply with the choice requirement, firms should provide consumers choice regarding uses of information that go beyond the reason for original collection (such as completing a transaction). Access requires that firms offer consumers reasonable access to their information and an opportunity to review and correct any errors. Security requires that firms take reasonable security measures to protect personal information.

B.3. FTC 2012 Privacy Guidelines

The latest FTC guidelines expand on these practices and articulate more substantive protections to satisfy instrumental objectives. "These best practices include making privacy the 'default setting' for commercial data practices and giving consumers greater control over the collection and use of their personal information through simplified choices and increased transparency. Implementing these best practices will enhance trust and stimulate commerce."[41] The principle embodying these objectives has been labeled "privacy by design" and dictates that companies include substantive protections into their practices, including data security, impose reasonable limits on data collection, implement comprehensive data management procedures (including personnel assessment and adequate oversight of third parties and service providers), and adopt sound data retention and accuracy practices.[42] It is important to note that, unlike disclosures, several of these protections might not be reflected in privacy policies.[43]

The "collection limitation" principle requires firms to limit data collection to purposes consistent with the context of the transaction or of the relationship between the firm and the consumer.[44] Whenever companies collect information beyond these contextual frameworks, firms are asked to disclose this to consumers at appropriate times. In addition, companies should implement data retention and disposable policies.[45]

The FTC also improved its notice requirement in 2012 by urging firms to adopt simplified notices and consumer choices and urging firms to "increase the transparency of their data practices."[46] Some of these privacy choices should be given outside the privacy policy itself. The report demands that "privacy notices should be clearer, shorter, and more standardized to enable better comprehension and comparison of privacy practices."[47] Finally, the "access" principle asks that companies provide reasonable access to the consumer information they maintain.

B.4. White House 2012 Privacy Bill of Rights

The principles in this document were created with the goals of being adopted into legislation and becoming templates for firm codes of conduct. The first principle,

---

[41] See supra note 12at i.

[42] Id. at vii.

[43] According to the report, adoption of these substantive protections would also allow firms to focus on more streamlined disclosures of their information practices. Id. at 23 ("By shifting burdens away from consumers and placing obligations on business to treat consumer data in a responsible manner, these principles should afford consumers basic privacy protections without forcing them to read long, incomprehensible privacy notices to learn and make choices about a company's privacy practices.").

[44] The FTC explained that "[i]n order to protect consumer privacy, there must be some reasonable limit on the collection of consumer data. General statements in privacy policies, however, are not an appropriate tool to ensure such a limit because companies have an incentive to make vague promises that would permit them to do virtually anything with consumer data." Id. at 27.

[45] Id. at 28. ("The commission confirms its conclusion that companies should implement reasonable restrictions on the retention of data and should dispose of it once the data has outlived the legitimate business purpose for which it was collected.").

[46] Practices for commonly accepted uses of data, such as product fulfillment, legal compliance, fraud prevention do not require choice.

[47] Id. at 61. Disclosure and choices should be given for collection of sensitive information for first party marketing as well as for the collection of geolocation information. Id. at 47.

"individual control," asks that firms give consumers some control over their personal information. The choices given should be in clear ways at convenient times. Consumers should be given the option to withdraw or limit their consent to data collection and uses. The second, "transparency," asks that firms provide clear and meaningful descriptions of their privacy practices. The third, "respect for context," mimics the FTC 2012 collection limitation principle and asks firms to collect and use data in ways consistent with the context in which the information is shared. For uses that go beyond this, firms must offer heightened notice and choice. "Security" asks that firms adopt reasonable security measures to control risks of loss, unauthorized access, and improper disclosure.

The principle of "access and accuracy" asks that firms adopt reasonable measures to ensure data accuracy and grant access to consumers who wish to review, amend, or delete their data. "Focused collection" states that consumers have a right to limited collection of their information and asks the firms adopt procedures to securely dispose or de-identify data when they no longer need it, except for legal obligations to retain it. Finally, "accountability" requires that firms be accountable to enforcement authorities, adopt adequate privacy protection measures, including training employees, having contracts with third parties binding them to adhere to certain privacy standards.

B.5. US-EU Safe Harbor Agreement

This scheme creates a voluntary mechanism enabling US organizations to qualify as offering adequate protection for personal data transferred from the EU. To comply with its requirements, adherents must register yearly with the Department of Commerce and adopt publicly available privacy policies. The policies must include the following: offer notice by disclosing the type and uses of information collected and providing a way to contact the organization with questions. Like the latest FTC and WH guidelines, the notices must be presented clearly and conspicuously before consumers share their information. Consumers should also be given opt-out choices when information is shared with third parties or used for purposes different than those associated with the original collection. Opt-in choice should be given in dealing with sensitive information.

The principle of "onward transfer" requires firms that disclose personal information to third parties to certify that it had entered into agreements with them to take some degree of privacy protection for that information. Firms must also employ reasonable security measures to safeguard personal information. "Data integrity" requires firms to use the information for the purposes stated in its privacy policy. It also requires that consumers be given a right to access their data. Finally, self-certifying firms must identify independent recourse mechanisms that can address unresolved complaints.

B.6. A Theoretical Benchmark: Maximum Protection

The guidelines above tend to require disclosure of practices but, in many cases, do not specify whether those practices need to be "protective" of the consumer. I contrast the guidelines above against an extreme benchmark of consumer privacy protection based on what is, for each term, the most protective specification possible. For example, with respect to sharing data with third parties, the most protective position is simply to share no data at all, period.

Note that I am *not* claiming that a policy that matches this benchmark is providing the optimal policy, i.e., the one that maximizes economic gains for firms and consumers. Sharing may be desirable for consumers when it directs them to goods or services that they want, for example. And, if the firm makes a profit from this sharing, and passes it on to consumers in the form of lower prices, all the better. In this case, the maximum-protection benchmark is overly protective. Moreover, the idea of a maximum protection benchmark is far from the spirit of notice and choice regimes (but it is moderately aligned with the latest approaches that seek that firms incorporate additional substantive protections). Nonetheless, it provides an interesting "absolute" standard of comparison for the regulatory guidelines and offers a potentially useful comparison of current practices.

C. Additional Studies

Most of the consumer protection initiatives regarding consumer information privacy have focused on privacy policies, their content, and their ability to stimulate market forces. They are known to be long, confusing, and demand a college-level reading ability.[48] It has been estimated that the average individual would need 201 hours a year to read all the privacy policies of the sites she visited.[49] It is thus not surprising that people don't read them.[50]

Early evidence suggests that companies have also failed to adopt FIPs. A study by Mary Culnan reported that about 20% to 40% of 100 policies studied complied with FIPs outlined by the FTC.[51] The report employed a generous notion of compliance (i.e., the notice requirement would be met if the privacy policy disclosed at least some information collected from the individual, not all of it) and also found that firms did not adhere to voluntary codes, such as privacy seal certification programs, as much as originally hoped.

A 1999 study by EPIC tracked ten terms in the 100 most popular sites and found weak compliance with the preliminary version of the 2000 FTC guidelines.[52] A 2004 study commissioned by the EU found that firms only weakly complied with the standards of the US-EU Safe Harbor Agreement.[53] More recently, a study using a machine learning approach to examine the standardized privacy policies of thousands of financial

---

[48] Carlos Jensen & Colin Potts, Privacy Policies as Decision-Making Tools: an Evaluation of Online Privacy Notices, 6 SIGCHI 471 (2004); Xinguang Sheng & Lorrie Faith Cranor, Evaluation of the Effect of U.S. Financial Privacy Legislation Through the Analysis of Privacy Policies, 2 INFO. SCI. J. OF L AND POL'Y 943 (2005).

[49] McDonald & Cranor, supra note 26.

[50] Privacy Leadership Initiative. Privacy Notices Research Final Results (Nov. 2001), available at http://www.understandingprivacy.org/content/library/datasum.pdf.

[51] See Mary J. Culnan, Georgetown Internet Privacy Policy Survey: Report to the Federal Trade Commission (June, 1998), available at http://www.msb.edu/faculty/culnanm/gippshome.html.

[52] See Electronic Privacy Information Center, Surfer Beware III: Privacy Policies Without Privacy Protection (Dec. 1999), available at https://epic.org/reports/surfer-beware3.html; Chris Jay Hoofnagle, Privacy Self Regulation: A Decade of Disappointment (Mar. 4, 2005), available at http://epic.org/reports/decadedisappoint.html.

[53] Jan Shont, Maria Veronica Perez Ansinari & Yves Poullet, Safe Harbour Implementation Study (Apr. 19, 2004), available at http://ec.europa.eu/justice/data-protection/document/studies/files/safe-harbour-2004_en.pdf.

institutions found that very few comply with the Gramm-Leach-Bliley Act's mandated disclosure requirements.[54]

In 2010, Bamberger and Mulligan interviewed nine chief privacy officers from large firms to examine how corporations manage privacy and comply with current laws.[55] Interviewees claimed that their privacy practices were being driven by consumer demand, the FTC guidelines and threat of FTC enforcement actions, and state data breach notification statutes. While noting that some guidelines were ambiguous and hard to comply with, the respondents stressed the need for flexibility given the rapidly evolving nature of technology, the fluidity of the products and services offered, and the different ways in which information would be used. Respondents also stated that they they tended to look to the terms in the US-EU Safe Harbor Agreement for guidance: "[W]e end up defaulting to the highest common denominator […] which really right now is Europe, and enforcing a fairly European code of conduct when it comes to privacy and information protection." [56] The authors reported that respondents viewed compliance with the guidelines as a floor, or just a minimum.[57] The conclusion of this interview-based study was that firms had adopted privacy protections and were highly receptive to FTC guidance, thus suggesting that the current model was succeeding in creating desirable consumer information practices.

III. What's in a Privacy Policy? An Analysis of Content and Compliance

The disparate judgments of the effectiveness of the current regulatory model discussed above indicate the need for evidence that goes beyond small-sample studies or interviews. We now describe a large-sample, detailed empirical analysis of the content of privacy policies and their compliance with the most recent FTC guidelines and other relevant guidelines.

A. Sample

The 261 sample firms' policies are drawn from seven online markets where consumers often share personal or sensitive information: adult (17 firms), cloud computing (28), dating (40), gaming (20), news and reviews (18), social networks (89), and special interest message boards (49). These are markets where information sharing is relatively more salient than in others where information sharing is a secondary aspect of

[54] Lorrie Faith Cranor et al., Are They Actually Any Different? Examining Thousands of Financial Institutions' Privacy Policies (2013), available at http://www.econinfosec.org/archive/weis2013/papers/CranorWEIS2013.pdf; see also Michael Birnhack & Niva Elkin-Koren, Does Law Matter Online? Empirical Evidence on Privacy Law Compliance, 17 MICH. TELECOMM. TECH. L. REV. 337 (2011) (conducting an empirical examination of compliance with Israeli information privacy laws by examining the privacy practices of 1360 active websites and finding low levels of compliance).

[55] Kenneth A. Bamberger & Deirdre K. Mulligan, Privacy on the Books and on the Ground, 63 STAN. L. REV. 247 (2010)

[56] Id. at 270.

[57] *See* Bamberger and Mulligan at 266.

the particular transaction, such consumer retailers or news sites. There are interesting differences in the nature of privacy concerns across these markets. Individuals share more or different information on dating and social network sites than on gaming and reviews sites. Indeed, the services offered by dating and social network sites depend on information shared by users. Other markets involve activities that are highly private (in the case of adult sites), or involve significant losses in the event of, for example, a security lapse or equipment failure (in the case of cloud computing). I start with an analysis of the whole sample of policies and then turn to market differences in a subsequent section.

The firms involved do business in the United States but may also have overseas operations. They include giants like Facebook and Google, and many smaller firms like veggiedate.com. At the time of sample selection, there was no obvious single list of sites to include in this study; I chose sites from various publicly available lists. An initial collection of 150 firms came from an extensive list of relevant sites on Wikipedia in 2009, 2010, and 2011, times of sample selection.[58] In 2010, I became aware of www.100bestdatingsites.com and used it to increase the sample of dating sites by selecting those that catered to individuals in the United States. I obtained additional bulletin board sites from rankings.big-boards.com. Finally, I added 17 adult sites in 2015 based on Alexa traffic rankings (described below).[59] Although the data gathering process was somewhat piecemeal, I am not aware of any obvious selection bias. To investigate the representativeness of the resulting sample, I compared the firms in five of our markets against market share reports generated by IBISWorld, an industry research firm, and confirmed that the sample includes the top firms.

Table 1 summarizes company, service, and policy characteristics. About 4% of sample firms are nonprofits. These may have lower interest in sharing personal information. 27% of the sample firms are public, potentially a proxy for firm size and sophistication; however, none of the adult sites are associated with public companies. Firms who earn money from the service itself might have a decreased need to rely on the sharing of personal information as a source of revenue. 39% of sample firms offer at least a portion of their services for a fee, but there are differences across markets. 93% of dating sites, a little over half of cloud computing and gaming sites, and a quarter of all adult sites are on a subscription basis. The remaining markets do not offer subscriptions but offer premium access or the ability to purchase items for a price. These include 16% of social networks, 31% of message boards, and 28% of news and reviews sites. This last number includes firms like Amazon.com, who have review forums but also sell merchandise. Firms who earn money from the service itself might have a decreased need to rely on the sharing of personal information as a source of revenue. Alternatively, it may be a fundamental part of their broader business model.

Alexa ranking is a ranking of web sites from alexa.com and is based on the number of monthly visitors.[60] This, too, is a proxy for firm size and reputation, as well as

[58]The list was originally available at http://en.wikipedia.org/wiki/List_of_social_networking_websites.
[59] I targeted 20, a round but manageable number, but three sites did not have policies.
[60] Alexa rankings offer approximate estimates of web traffic because they rely on the metrics provided by those users who install the Alexa Toolbar, which might not be representative of all Internet users.

the potential volume of privacy-related information flow. A lower number indicates a more popular site; Google's Alexa rank is 1. The mean Alexa rank is 949,099 and the standard deviation is 3,766,538, indicating the large range in the popularity of sample firms. One of the self-regulatory efforts has been to encourage firms to adopt codes of conduct in the form of privacy seals, which are associated with a standardized set of privacy practices. I track whether a particular firm affirms in its privacy policy that it adheres to a particular privacy seal, such as TRUSTe, or claim to adhere to a particular framework to comply with the laws of foreign states, such as the US-EU Safe Harbor.[61] Thirty percent of the sample firms claim certification by at least one privacy seal or conformity with a self-harbor agreement. This low take-up rate is consistent with prior evidence.[62] More important, this might translate into a rather limited role of third party information intermediaries.

There are wide differences across markets, however. Sixty eight percent of cloud computing firms claim some type of certification. Companies in this market might feel more pressure to signal their commitment to privacy protections. Consumers' potential losses associated with the loss or leaking of uploaded information are likely large and might demand such protections, and corporate clients of these firms may insist on certain certification measures guaranteeing some acceptable levels of security. Note that not a single adult site claims a certification.

The sample firms were collected over years, but the privacy policies and other variables used in this study were all collected in June 2013, with the exception of the adult sites' policies, which were collected in August 2015. Most firms list the date of last update explicitly on their policies. For a few dozen more firms I were able to measure the year of last update using aspects of the longitudinal data set described in Marotta-Wurgler and Svirsky (2015).[63] But for a few dozen firms I have no method of estimating the year of last update. On average, contracts in force in June 2013 were last updated in 2011 (median 2012). Cloud computing firms have slightly recently more updated contracts, perhaps because this is an emerging market. The year of last update will help shed light on the extent to which companies respond to guideline changes.

There has been a recent push by the 2012 FTC and White House guidelines to standardize and shorten contracts, including measures being considered in California[64],

---

Nonetheless, Alexa is the industry standard, and the rankings appear sensible—Google is the most popular site, Facebook is number two, etc.

[61] Of the 78 firms claiming to adhere to a benchmark or seal (unreported), 58 claim to adhere to the US-EU Safe Harbor and 25 to the US-Swiss Safe Harbor, which enable firms with a presence in the EU and Switzerland to comply with the EU laws' heightened standard. Sample companies claim to adhere to 14 other certification programs in the EU and Australia (such as United Kingdom Information Commissioner's Office and the Australian Best Practice Guidelines for Online Behavioral Advertising).

[62] *See* Joel Reidenberg, *E-Commerce and Trans-Atlantic Privacy*, 38 HOUS. L. REV. 38 (2001).

[63] *See* Florencia Marotta-Wurgler & Daniel Svirsky, *Who's Afraid of the FTC? An Analysis of the Deterrent Effect of "Unfair and Deceptive" Practices using Privacy Policies* (Working Paper, 2015) (using a dynamic panel comprised of weekly snapshots of the sample firms' privacy policies from 2009 until 2014).

[64] *See* A.B. 242, 2013 Leg. Reg. Session (Cal. 2013) (proposing to amend California Online Privacy Protection Act to "require the privacy policy of a commercial Web site or online service to be no more than 100 words, be written in clear and concise language, be written at no greater than an 8th grade reading level.").

but existing guidelines prescribe no limit to contract length. The average length in our sample is 2,176 words long, approximately the average length of End User Software License Agreements.[65] Adult sites have the shortest contracts, with an average of 1,356 words—less than half the length of gaming sites' policies. The analysis will show that much of this difference is explained by adult sites' tendency to claim little collection and sharing.

Table 2 shows correlations. I transform the Alexa ranking into a "popularity" measure by taking the negative of the log of the Alexa ranking; this yields a more symmetric distribution and a more intuitive definition whereby higher popularity simply means more traffic. Public firms tend to be associated with more popular websites and their privacy policies are more likely to claim certifications and be longer and more recently updated. This is not surprising given the presence of in-house counsel and the pressures of reputational sanctions and likely stock price punishment for major privacy violations. Finally, nonprofits have somewhat shorter policies.

B. Contract Substance and Compliance with the Guidelines

This sub-section examines the content of privacy policies and evaluates the relative degree of compliance with the self-regulatory guidelines to help assess the desirability of the current regime. I track 49 terms across seven categories: Notice, Sharing, User Control, Security, Data Practices, Enforcement, and Privacy by Design. Why 49 terms? This is the number of terms discussed by at least one of the five guidelines that I track. The division of terms into categories is more subjective (they tend to match the structure of privacy policies), but this is less important than the terms themselves.

Each contract was read and terms codified by hand. Contracts were divided and assigned to two of eight law students. To increase accuracy, each member of a pair read the entire contract and graded a specific portion of it independently. Each contract was thus graded twice. I revised their coding periodically.

Over weekly meetings, we discussed any discrepancies and collectively decided on the proper classification.[66] These discussions could be long, because contracts often include ambiguous clauses and gave rights that cannot be exercised. For example, consumers are commonly told that they are given a choice as to how their personal information can be shared, but do not explain how that choice can be exercised and do not appear to offer the choice outside the contract.[67] Or, firms would state that their information would not be shared, only to list several lines indicating that it might be shared with third parties, or with user consent—but it was unclear whether the choice was

---

[65] Florencia Marotta-Wurgler, *What's in a Standard Form Contract? An Empirical Analysis of Software License Agreements*. 4 J. EMPIRICAL LEGAL STUD. 677 (2007).

[66] I used Cohen's Kappa to measure intergrader discrepancies. This method accounts for chance disagreements across readers. Kappa began at 0.64 after the first round of grading and rose to 0.88 after group discussion of ambiguities in the text. The author settled remaining disagreements.

[67] The FTC settled a case in 2015 under Section 5 of the FTC Act against Nomi Technologies for engaging in misleading practices such as these, where the firm claimed in its privacy policy to give certain opt-out options to consumers but failed to do so in practice. *In re* Nomi Technologies, Inc., File No. 132 3251 (FTC, Sept. 3, 2015), *available at* https://www.ftc.gov/system/files/documents/cases/150423nomiorder.pdf.

opt-in, opt-out, or a choice at all. It is impossible to know whether these situations are the result of careless drafting or deliberate obfuscation.

The legal interpretation of such clauses is complex. Many appear to be deceptive and thus not enforceable. Furthermore, if the statement such as "we do not share your information" is considered a warranty, then a company cannot later disclaim it in the same document. Our heuristic was that choices that were not clear were not counted as choices and statements that were later retracted were not counted as affirmations of fact. Even though courts would most likely interpret such retraction against the drafter, I did not code them as choices because the purpose is to document stated practices.

The terms are divided into seven categories of related terms that are loosely based on the guidelines, but which more closely follow the content and order of privacy policies as they are written. Thus, they should provide a fairly comprehensive snapshot of what one would expect in a privacy policy.

B.1. Notice

The main analysis of privacy policy content presented in this paper is in Table 3. I start with an extensive set of terms (21 terms) involving notice. For notice to effectively inform consumers the terms in the policies must be clear and privacy practices must be spelled out completely, especially since there no clear default rules that can fill in gaps in the face of contractual silence are unclear.[68]

The first few terms describe the extent to which notice of the policy's existence and content is prominent (as encouraged by the FTC 2012 and White House guidelines[69]). While most companies make their privacy policies available somewhere on their home page (N1), only 19% require users to expressly agree to them before proceeding (N2). While this goes against the directive to make policies salient, this distinction may not matter too much in light of the evidence from software license agreements that clickwraps tend to be ignored.[70] I also measure the extent to which firms embraced the FTC's 2012 recommendation that privacy policies include short, or "layered" notices, summarizing the most important terms of the policies (N3). The practice has yet to pick up, as only 22% of contracts include a short notice. A concern is that firms may not have had time to adopt to the new guidelines, but note that a preliminary report with the FTC recommendations had been circulating since 2010. I address this further in Section IV.

---

[68] Until recently, the FTC has brought actions mostly against firms that violated explicit statements in their privacy policies, leaving room to conclude that undisclosed information practices are allowed unless they are unfair and deceptive in other ways, or violate other state information privacy laws. Yet a number of FTC actions and state law proposals have complicated this understanding. For example, in 2012, the FTC brought an action against Epic Marketplace for accessing the browser history of its users to deliver targeted advertising. The FTC found that Epic's failure to disclose this practice in its privacy policy was a material omission that violated Section 5 of the FTC Act. *In re* Epic Marketplace, LLC and Epic Media Group, File No. 112 3182 (FTC, Dec. 5, 2012), *available at* https://www.ftc.gov/sites/default/files/documents/cases/2012/12/121205epicorder.pdf.

[69] *See supra* note 34 and accompanying text.

[70] *See* Yannis Bakos, Florencia Marotta-Wurgler, & David Trossen, *Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts*, 43 J. LEGAL STUD. 1 (2014).

The next series of terms in this category report the extent to which firms disclose which types of data they collect. Collection is pervasive. The overwhelming majority reports that they collect contact, computer (IP address, etc.) and interactive (browsing behavior, search history, etc.) information. Only a small percentage of firms state that they do not collect contact (4%), computer (3%), interactive (13%), financial (23%), and content (19%) information (N4-N8). 41% state that they do not collect sensitive information (N9), perhaps due to the influence of the EU Data Privacy directive, which treats sensitive information differently. In addition, the FTC guidelines also ask the firms give users notice and choice regarding the collection of sensitive information.

A fairly large fraction of firms simply fails to disclose their collection practices. As noted, silence is problematic because it is unclear whether firms are allowed to engage in undisclosed behaviors. Moreover, a number of terms reveal that most firms have failed to impose constraints on some of their most critical data collection practices, such as collecting personally identifiable information (PII) for only internal (N12) or context specific purposes (N13), as encouraged by the latest FTC and White House guidelines.[71]

Like collection, sharing is also extensive. Sixty-two percent allow third parties to track user behavior (N15) and only 14% explicitly state that they do not allow third party tracking. The remaining firms do not disclose their practice. Only 9% of firms identify recipients of sold or shared data (N16) and only 7% define words such as "affiliates" or "third parties" when they use them (N17). The latter might be due to the dynamic nature of this business, where parties and uses of data are constantly arising.

The issue of notice is one where the maximum protection benchmark diverges from regulatory guidelines. This divergence is in the nature of the notice and choice regime, as *any* disclosure, even if communicating an invasive practice would be deemed to comply with the standard.

B.2. Sharing

I track eight terms that measure information sharing practices. While consumers know some of the types of data they have made available to a given website without reading it in a privacy policy, they cannot monitor the extent to which it is passed on to third parties, or the behavior or the security practices of third parties with respect to their personal information. Most they have to go on is the contents of the privacy policy.[72]

Only 13% of sample firms state that the affiliates and subsidiaries with which they share information (SH1), and 20% of contractors, are bound by their privacy policy (SH2). The typical combination of not naming third parties and not binding them to the same policy thus leaves consumers completely in the dark about the uses of their data, as advised by the guidelines.[73] Only 8% of firms report that they have a contract with third parties establishing how disclosed data may be used (SH7).

---

[71] For a detailed account of the role of context, see HELEN NISSENBAUM, PRIVACY IN CONTEXT (Stanford Law Books ed., 2010).

[72] Alternatively, consumers might be able to install plug-ins in their browsers that could identify first and third parties with whom each site shares data.

[73] Of course, naming dozens of third parties in the privacy policy would likely do little to improve consumers' understanding of firms' privacy practices.

A number of guidelines have encouraged firms give consumers choice and control over the sharing of their information. But choice is not widespread. Only 26% of firms ask consumers to opt-in to share information (SH8), the most consumer friendly mechanism and recommended by current FTC guidelines, and 7% offer at least an opt-out mechanism. Other firms do not offer an option (or the question is inapplicable). For the most part, firms assume little responsibility for the data that they share. These numbers reveal that consumers would learn little about firms' information sharing practices and that substantive protections are weak.

B.3. User Control

I track four terms pertaining to user choice and control of data, such as having the ability to access and correct any information. Fifty seven percent of firms give users the ability to adjust their privacy settings (UC1). 72% allow users to access and, usually, correct their personal data. This fairly high percentage is probably not representative of the larger population of privacy policies, because many of the firms in markets selected for this study allow users to revise their information on an ongoing basis due to the nature of the services offered (e.g. dating sites, social networks). While 56% of firms claim that user data can eventually be deleted or anonymized (UC3), the ashleymadison.com hacking of early 2015 is a reminder that firms do not always follow such requests (and, more generally, a reminder that this methodology allows me to measure only what firms claim to do, not what they do). Only 2% of firms offer consumers a choice about what happens to their personal information if the company is sold or goes bankrupt (UC4).

B.4. Security

The next seven terms relate to measures undertaken to protect data accuracy and security. Despite the increased attention to security breaches and the costs associated with them, most polices do not address data security in a complete way, even though they might have implemented reasonable practices that are not observable from the privacy policy. One possible reason for this might be a desire to avoid liability for breach of contract, while at the same time reducing the probability of data breaches by taking reasonable precautions. Thirty two percent of firms claim to have reasonable procedures to ensure data accuracy (SEC2). Only 2% guarantee accuracy (SEC1), which would seem to be a difficult guarantee to ask, but this guarantee is written into all five sets of guidelines.

As expected, most firms reserve the right to disclose personal information to comply with the law, protect a crime, or defend its own rights (SEC3, SEC4). 46% of sample firms identify specific security technologies like encryption (SEC 7) and 45% describe other protections in their operating procedures such as restrictions on the set of employees with access to data (SEC6).

B.5. Data Practices

Some of the most substantive recommendations by the FTC involve data practices, which are related to security. I track three terms to understand certain data practices. The

related guidelines are broadly ignored. Only 6% of sites state the period of data retention (DP1). Most do not state what happens to personal data when the account is closed (DP2), and only 1% states their procedure for disposing of unused data (DP3).

B.6. Enforcement and Others

The final class of terms relates to enforcement. The vast majority, 94%, includes contact information for privacy questions or concerns (E1). Current FTC guidelines discourage disclaimers of liability for security failures, but 58% of firms disclaim said liability (E2). Thirty percent of firms claim a privacy seal or certification (E4). Of these, only 24 (9% of all firms) claim a seal other than the Safe Harbor or international compliance agreement. To anticipate results presented later, even firms that state that they adhere to US-EU Safe Harbor guidelines do so only to a limited extent.

B.7. Privacy by Design

The last two terms measure whether firms adopt "privacy by design" measures. Again, this is an attempt to encourage firms to adopt more substantive measures. They generally do not. Only 13% state that they conduct periodic compliance reviews of data and security measures (PBD1). A reasonable 43% of cloud computing firms, however, claim to do so. Also, only 5% of sample firms (and 29% of cloud computing firms) contain self-reporting measures in case of privacy violations (PBD2).

C. A Summary of Compliance

The terms I measure do not always fit the guidelines of the principles perfectly. Some are not directly mentioned in the guidelines but, in my opinion, closely approximate a stated practice or objective. Also, some terms included in some principles are likely to be unaccounted. On some points, the guidelines are as nuanced or vague as some of the privacy policies they address. Still, I believe the data are capable of identifying broad differences in substance and compliance across principles and terms.

At the term level, Table 3 suggests that compliance is often higher on the "easiest terms": disclosures of the collection of information that was already obvious to the user because he or she had entered it, such as financial information, or disclosures of low probability events, such as using your identity for advertising. Compliance with terms that are likely to require firms to incur costs or provide substantive protections is less pervasive. Examples here include contracts with third parties to guarantee data use and sharing limitations, data accuracy guarantees, or data retention periods. In other words, there is more compliance with terms that are less protective of consumer information and less compliance with terms that are more protective.

Table 4 shows a summary of the number of contracts that comply with each of the guidelines, laws, and safe harbor agreements. (I omit a tabulation for the benchmark of maximal consumer protection because it is not clear what overall level is desirable. For the five sets of guidelines, the regulators, at least, prefer more compliance to less.) For example, the latest FTC guidelines are comprised of 27 out of the 49 terms that I track. The results show that just 1 of 261 privacy policies in the sample (the policy of an adult

site) complies with at least 80% of the terms comprising this standard. The modal degree of term compliance is 30-39%. As a matter of terminology, the FTC's 2000 Report to Congress referred to 41% compliance as "low."[74] This is in striking contrast to the conclusions from the survey responses given to Bamberger and Mulligan, where respondents claimed that SHA compliance was vital and posed a floor pushing firms to adopt more privacy protections.

Not a single contract complies with 90% or more terms regardless of the standard. (The same adult site policy is at the top of every distribution, reflecting a degree of consistency in guidelines.) It is particularly striking that 41 out of 261 policies comply with less than 10% of the 2000 US-EU Safe Harbor guidelines—just two, one, or zero of its 19 total terms.

Compliance with FTC guidelines is of particular interest, given that the sample firms operate in the United States and were asked to engage in self-regulation embracing these guidelines. Self-regulation might have become more effective over time as more firms adopted privacy policies and FTC enforcement actions became more pervasive and gave clearer indications of what might constitute problematic practices. But the results indicate that firms have become somewhat less likely to comply with FTC recommendations as they changed. Whereas 192 of all firms comply with 50% or more of the FTC 2000 guidelines, only 66 have, to date, achieved the same level of compliance for the 2012 guidelines. One explanation is that the 2012 guidelines add more substantive protections, making it harder and costlier to comply.

The US-EU Safe Harbor has arguably more teeth because it requires firms that claim to adhere to it to comply or face FTC enforcement actions with a higher degree of certainty, given that what constitutes a violation of it is fairly clear. The results show that even with a clearer threat of enforcement actions, companies don't abide to the Safe Harbor as strongly as has been recently stated. The numbers in brackets show the compliance fraction for firms that claim to adhere to the US-EU Safe Harbor. Only 12 firms who claim compliance actually appear to comply with at least half of the terms that comprise the safe harbor. The remaining firms have lower compliance rates. Hence, most of the companies that claim to adhere to the Safe Harbor, including some of the most popular websites, are listing practices that violate large portions of it.

The FTC has brought a number of actions (especially in the past two years) against firms for failing to comply with the US-EU Safe Harbor requirements. But all of these resulted in settlement agreements where such firms were prohibited from further misrepresenting compliance with the terms of the Safe Harbor. Given the lack of sanctions and relative low probability of being the subject of such action, the low level of compliance is perhaps not surprising.[75] In addition, a decision by the European Court of Justice in October of 2015, Schrems v. Data Protection Commissioner, declared the US-

[74] *See* FTC, *Privacy Online: Fair Information Practices in the Electronic Marketplace: A Report to Congress* (May, 2000), *available at* https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-fair-information-practices-electronic-marketplace-federal-trade-commission-report/privacy2000.pdf.

[75] *See, e.g.*, Press Release, FTC, FTC Settles with Twelve Companies Falsely Claiming to Comply with International Safe Harbor Privacy Framework (Jan. 21, 2014), *available at* https://www.ftc.gov/news-events/press-releases/2014/01/ftc-settles-twelve-companies-falsely-claiming-comply.

EU Safe Harbor Agreement invalid because it failed to provide adequate protection to the data of citizens of the EU given US intelligence activities.[76]

Based on the analysis in Table 3 and 4, it is hard to escape the conclusion that self-regulatory guidelines, even with FTC enforcement actions, are not shaping modern privacy policies all that much, regardless of views on the contrary. The degree of compliance is modest by any measure. It is true that the terms I measure are those that generally appear in privacy policies and do not always fit the principles of all guidelines perfectly. Some are not directly mentioned in the guidelines but, in our opinion, closely approximate a stated practice or objective. Also, some terms included in some principles are likely to be uncounted. On some points, the guidelines are as nuanced or vague as some of the privacy policies they address. Nonetheless, I believe the data are capable of identifying broad differences in compliance across principles and terms.

It is usually difficult to make strong statements about the optimality of contracts, standard form or otherwise. Optimality generally depends on unobservable preferences of consumers and costs and revenues to firms. For example, a good may come with no warranty, but one cannot rule out the possibility that consumers prefer a slight discount to a warranty, even a long one. But in the case of privacy policies it is clear that the current situation is unlikely to be producing optimal policies. Absent other mechanisms, "competing on privacy" cannot work if, as is often the case, the consumer simply cannot figure out the firm's policies, and there are no default rules akin to UCC's Article 2 to serve as a backstop.

IV. Differences Across Markets, Within Markets, and Guideline Changes

One goes to a sushi restaurant without fear because one expects the food to be unspoiled. Sushi restaurants keep their inventory frozen in order to kill bacteria and ensure that fish caught days ago and oceans away may be served "fresh." The fact that the Department of Health also enforces rules about freezing fish may thus not be necessary. Compliance with regulation might be incidental, not deliberate; as a result of consumer demand for unspoiled sushi—"market forces"—it would likely happen anyway.

To shed additional light on how effective the current notice and choice regime is in this context—specifically, the FTC's 2012 guidelines—it is important to deal with this sort of identification problem. How much of the compliance that one *does* observe should be attributed to the guidelines themselves? To what extent does compliance translate into substantive protection? I address these questions here. A tentative conclusion will be that at least some portion of observed compliance is incidental to regulatory guidelines. In this sense, the prevailing regulatory model may be even less influential than it appears. At the same time, market forces may be, at least to a modest extent, filling in the gap.

A. Differences Across Markets

Some firms comply with guidelines more than others. One can learn more about deliberate versus incidental compliance by studying these cross-sectional differences. Do

[76] *See* note 10 .

they match what one would intuitively expect given the nature of the markets and services involved?

An obvious way to split the data is by market. The top panel of Table 5 shows the average rates of compliance with FTC 2012 guidelines by market and term category, and the bottom panel shows an average measure of protection, which is akin to "compliance" with the maximally protective set of terms set out in Table 3.

"Compliance" is measured simply as the fraction of terms in that category that satisfy the guidelines. The first column shows average compliance for the full sample of firms and terms. The top left number, an OVERALL compliance measure of 0.39, indicates that the average term is complied with by 39% of firms. Put another way, the average firm complies with 39% of the terms in the FTC 2012 guidelines, corresponding to between 10 and 11 out of the 27 guidelines.

"Protection" differs from the compliance measure in two ways. The first is that it takes into account the full set of 49 terms that are mentioned in any of the guidelines. The second is that it is slightly more sophisticated in not assuming that all terms are equally important. The protection index weights terms according to the degree of attention paid to it across standards. For example, SEC1 (data accuracy guarantee) is a component of all five standards, whereas SEC5 (user alerted of government requests for her information) is a part of only two. The former term is then given 2.5 times as much weight in this protection measure. Comparing the compliance index to the protection index allows us to take a rough look at whether the pattern of compliance that we see represents substantive protections for the consumer (when compliance does occur) or a more technical type of compliance with idiosyncratic, less important terms. Nonetheless, partly because the FTC 2012 requirements sometimes involve substantial protections, the measures are highly correlated (0.88). Of course, weights would ideally take into account consumer preferences, but absent data on this, my approach may offer an approximation.Figure 1 shows the distribution of overall compliance for the full sample and by market. A reference line at 50% compliance is provided. This is a more granular version of the fourth column of Table 4. The rest of the figure shows histograms by market. The comparative superiority of adult and cloud computing sites, in terms of overall compliance, is apparent from these histograms.

Cells in Table 5 are colored green when policies from that market are more compliant (or protective) than average to a statistically significant degree. Cells are colored red when policies are significantly less compliant than average.[77] The differences across markets are striking. For starters, the patterns of green and red in overall protection index essentially match those from the overall compliance index, suggesting that the latter are capturing robust differences in actual substantive protections.

For adult sites, more than two-thirds of notice terms are compliant with FTC 2012 guidelines, but in no other market is the rate of compliance more than half. In other words, despite being several hundred words shorter than average, adult sites' policies provide considerably more detailed notice of many critical privacy practices, and this translates to substantive protection for the consumer, as disclosures reveal little collection.

[77] As an aside on statistical power, Table 5 suggests that the sample size is sufficient to identify meaningful market differences. In almost all cases, a difference of 10% from the average is flagged as significantly different. In some cells, even differences as unimportant as 3% are identified as statistically significant.

Adult sites are even more exceptional in terms of sharing practices. More than two-thirds of their sharing-related terms are compliant, while in other markets the rate of compliance varies from 19% to 43%. More important, this translates into marked differences in the level of protection. Table 3 indicates that to be compliant with sharing terms tends to mean limited sharing of information, and this is the case with adult sites. It is obvious that consumers of adult sites would prefer this and arguably demand. If there is a market where privacy is especially salient to consumers, it is this one.

Contrast this with dating, gaming, or social network sites. In these cases, the functionality of the site depends on connecting people with similar preferences. The collection and sharing of information with other users of the platform, or information that ties the user to that particular service, would be expected. Collection and sharing of such information with unknown third parties, however, might not be as salient to consumers, thus weakening any market demands for increased protections.[78] In the case of gaming, liberal sharing practices are driving down its overall performance to a statistically significantly below average level.

Adult sites appear to take a less pro-privacy stance in allowing users to control privacy settings or other aspects of their information. This is somewhat unfair, however, because they generally collect and share little data in the first place. There is usually little data for the user to control. Message boards tend to limit the ability of consumers to delete or anonymize their information. This contributes to their significantly lower rate of compliance and protections here. The higher degree of compliance by dating sites presumably reflects the need to ensure current and correct information on personal characteristics, relationship status, and the like.

One of the most noteworthy results in the remainder of the table involve cloud computing. Cloud computing policies comply with FTC data security and security-related guidelines to a far greater extent than other markets. This is intuitive. The security of data storage is a fundamental expectation of cloud computing users. Consider the possible economic losses that could result from security breaches.

In particular, cloud computing sites are more explicit about security measures and, as mentioned earlier, are much more likely to claim compliance with third-party certifications. The relevant privacy by design guideline included in the FTC 2012 guidelines, to require periodic reviews of data security measures, is widely ignored—except by cloud computing firms. Also, the only enforcement-related requirement for compliance with FTC 2012 guidelines is not to disclaim liability for security measures. Cloud computing policies are less likely to disclaim liability and this is feasible given their stronger security precautions in general.

A last interesting result again involves the low enforcement-related protections of adult sites. Again, this is to some extent misleading, because they collect little information in the first place. There aren't many collection or sharing provisions to enforce, so there is no reason to disclaim liability for data leaks, for example.

---

[78] *See* Fred Stutzman, Ralph Gross & Alessandro Acquisti, *Silent Listeners: The Evolution of Privacy Disclosure on Facebook*, 4 J. PRIVACY & CONFIDENTIALITY 7 (2012) (finding that Facebook users shared more information with friends on the platform as they increased and customized protections their privacy settings, leading the authors to conclude that users were unaware of third party disclosures).

Finally, an important pattern that does not appear in the table is the fact that *not even one* adult site claims compliance with a privacy seal or third-party certification, whereas the overall rate of adoption of seals and certifications in the rest of the sample exceeds 30%. Given the relatively protective nature of adult sites' privacy policies, it appears that the use of privacy seals is not strictly necessary to signal a certain level of commitment to information privacy, as it may be salient enough.

B. Differences Within Markets

The most natural interpretation of the differences across markets is that market forces beyond the guidelines themselves might be shaping privacy policies. Table 6 shows regressions of rates of compliance on product and contract characteristics. The regressions include market fixed effects, so the coefficients in the table reflect the within-market differences.

There are relatively few significant differences here, and, for the most part, the results that are statistically significant are similarly so for the compliance as well as the protection measures. One of the larger effects in the table is nonprofits' higher rate of compliance on sharing terms. Presumably, such sites have no business pressure to sell information to other firms. There are only scattered relationships between compliance measures and public versus private status, paid versus unpaid products, and site popularity. Note that almost all markets have free versions of the products (some more than, others), and that the results hold when controlling for this. The strong links between certifications and actual compliance are not surprising. It means that policies that claim certification are consistent with at least a portion of what the certification requires.[79]

Interestingly, policies that have been updated more recently are no more likely to comply with current FTC guidelines than those last updated years earlier. Recently updated policies are actually slightly less likely to comply with FTC guidelines: Those last updated in 2014 are, on average, compliant with 10% fewer of the 2012 guidelines (between two and three of its 27 terms) than policies updated in 2004 (-0.010*10=-0.10). The same pattern is true in all but one term category. I will return to this pattern in the third test, below.

Finally, longer contracts in the same market are generally more compliant than shorter ones, except with respect to sharing and particularly enforcement provisions. (This is true in notice terms as well, but it does not contradict the fact that adult sites score high on notice compliance despite their short policies because the regressions include market fixed effects.) As mentioned before, in the FTC 2012 guidelines, the only enforcement-related suggestion is not to disclaim liability for security measures. Apparently, the longer the contract, the more likely this disclaimer is to appear, all else equal. The pattern in overall compliance does not extent to the overall protection measure which takes account of more terms and attempts to importance-weight them.

For parsimony I do not report the market fixed effects, but statistically significant cross-market differences remain after controlling for these factors. The differences remain roughly as large as those in Table 5 and can account for at least half of the adjusted R-squared in the overall compliance and protection measures (unreported).

[79] Note that it is only a small portion from Table 4, second column, distribution in brackets.

C. Policy Changes Around Guideline Changes

The replacement of the FTC's 2000 guidelines with a new set in 2012 offers a potentially clean natural experiment to study deliberate compliance. I could simply compare policies in place before and after this event and reasonably attribute changes that match the new guidelines to the new guidelines themselves as opposed to a sudden shift in market demands for privacy.

The sample policies were collected in June 2013 and June 2015, and are not longitudinal, however, so the ideal test is not possible. Nonetheless, the date of last update is available for most of the policies. It varies from 2004 to 2015, thus surrounding the date of the introduction of the new guidelines. This allows for an interesting albeit imperfect approximation to the ideal test of deliberate compliance.

Of course, the fact that some of the policies in place in June 2013 had not been updated since 2004 already tells us that any response was at best incomplete. But one could learn more by studying policies at the term level. I divide terms into three groups—those unique to the 2000 guidelines, those unique to the 2012 guidelines, and those in both guidelines—and examine whether policies that were updated after 2012 are more likely to comply with the new guidelines than the now-immaterial ones. To the extent that at least a subset of firms deliberately updated in response to the guidelines, this should be the case.

Table 7 shows the results of regressions where the dependent variable is the fraction of compliance with the new, old, and overlapping terms, and independent variables include a dummy for post-2012-updated policies as well as control variables and market fixed effects, for which coefficients are suppressed because contain no surprises. To avoid the transition period, I exclude policies last revised in 2012. I should also remind the reader that a preliminary report with the guidelines was circulated in 2010 (excluding policies whose last update was in 2011 or 2010 does not change conclusions below).

The results suggest a muted response to the new guidelines. Policies updated more recently are, on average, no more likely to comply with the new terms than the old ones or the unchanged ones. The point estimates indicate that they are somewhat less compliant across the board, but none is statistically significant.

The experiment is imperfect because I do not know what the post-2012-updated subsample looked like before the change. But one can speculate about the likely direction of the bias. One possibility goes as follows. If the firms that *did* update were reacting to the 2012 changes, then one would have expected that the reason they did so, while others did not, is that they found themselves "too far behind" the new terms. It is natural to expect that if they took the time and expense of updating the policy, they would want to do so such that they were in reasonable compliance with the new terms. That is, one would expect a bias toward a positive coefficient on the key dummy. If there is such a bias, it was not meaningful enough to separate themselves from those that chose not to update. In fact, the point estimates suggest they did not even catch up.

D. Robustness

The main results—in particular, the observed pattern of market effects—appear robust to a number of alternative methodological choices. First, I have studied a more balanced sample in which I restrict attention to the twenty most-visited sites in each market. There are no material differences in the findings from the full sample. Despite losing half the data, the same market effects appear and remain statistically significant. This also means that small and obscure sites do not drive the results.

Second, I use least-squares regressions in Tables 6 and 7 for simplicity in understanding the magnitudes of the regression coefficients. The use of a Tobit model, which is more strictly appropriate in the case of a dependent variable bounded between zero and one, leads to identical inferences. This is not particularly surprising since, as Figure 1 indicates, the dependent variable is roughly bell-shaped and not a single data point hits 0.0 or 1.0.

V. Conclusion and Implications

For decades, the dominant regulatory approach to consumer information privacy has been the "notice and choice" self-regulation model. The notice and choice model emphasizes disclosure and encourages firms to adopt of substantive protections via self-regulation as opposed to regulatory supervision and enforcement.

The regime is in some ways an intermediate model between an extreme of no regulation or guidance at all, in which markets are left to do what they may, and a strict regime of detailed and enforced regulations. Its appeal is understandable. Relative to a regime of substantive regulation and strict enforcement, notice and choice is better able to accommodate the evolving technological environment and does not run the risk of trampling on true consumer and firm preferences. A regime that relies on market forces alone must also face the reality that the vast majority of consumers simply do not read privacy policies and thus cannot know the full implications of using the site; a modest degree of guidance and enforcement seems preferable to none. But evidence on actual firm practices and conformity to self-regulatory guidelines is scant.

I provide the first large-sample, comprehensive analysis of the actual content of privacy policies. The results shed light on current privacy practices and the extent to which they conform to self-regulatory guidelines. Specifically, I review the privacy policies of 261 firms in seven markets where information sharing is relatively more salient and privacy concerns are relatively significant, including adult sites, social networks and cloud computing. For each policy, I track the presence or absence of 49 terms pertaining to notice, information collection and sharing, data security, and other practices. I then step back and compare stated practices with various self-regulatory guidelines. I also study differences in compliance across firms and markets.

The analysis uncovers many specific facts about current privacy practices, but an overall summary is that compliance with current guidelines seems low. In terms of substance, data collection and sharing is widespread and difficult for the user to control. In fact, data collection and sharing practices are not just hard to control, they are hard or simply impossible to learn. Policies are long, complex, and often incomplete or silent on required dimensions. In general, compliance is lower for terms that require firms to offer costly or substantive protections or that place limits on their use of information collected.

An implication of this complexity is that it is difficult for intermediaries to simplify the terms in current policies in a consumer-friendly way or convert them into machine-readable policies able to be standardized or personalized.[80] Efforts by the latest FTC guidelines and the White House to simplify and standardize have been mostly ignored, even by policies that have been updated after these guidelines were publicized. Presumably, the incentive to comply is not high given weak enforcement mechanisms.

Compliance with the US-EU Safe Harbor Agreement is also low. In fact, policies that claim compliance with the agreement very often omit or contradict its requirements. While the recent *Schrems* decision by the European Court of Justice declared the SH invalid, our findings contrast with FTC assertions that losing the agreement will hurt consumers because firms that claim to adhere to it are taking compliance seriously.[81]

In addition to the generally low levels of compliance with guidelines, at least a portion of the compliance that is indeed observable seems likely to be reflecting market forces as opposed to deliberate efforts to meet guidelines. For example, adult and cloud computing sites stand out with more meaningful protections precisely in areas where would expect their users to care about—for adult sites, these are notice, data collection, and sharing; for cloud computing sites, these are data security and substantive data protections. This is remains the case even when controlling by whether the service is offered for free or payment. Some of protections these categories of sites offer go beyond those required by the latest FTC guidelines. The fact that these practices line up so intuitively with user preferences and market business models suggests that that some of the practices that we associate with "compliance" reflect more basic market forces which might produce similar outcomes anyway.[82] A plausible hypothesis is that this result could be the outcome of FTC focusing enforcement actions on the more compliant markets, yet this is not the case. Rather, FTC actions have focused on relatively large firms (like Google or Facebook) or on behavior that resulted in losses, such as data breaches.

One caveat of this study is that I can observe only stated policies, not behavior. Actual behavior does not always match stated practice; it may be more or less consumer-friendly. The FTC 2012 guidelines asked firms to take substantive protections that might not be reflected in the privacy policies, such as implementing privacy by design in product and service development. But regulators are in the same position as we are here. They must base their activities on what they can measure. Absent widespread random audits, the best they can do is measure what firms claim, and disclosure, in particular, has been an important component of notice and choice regulation for decades. FTC enforcement actions have focused on the statements made in privacy policies, highlighting policies as a core aspect of regulation.

In conclusion, and in contrast to assertions by some commentators, the current regulatory model appears to be having a limited impact on privacy policies. By replacing media anecdotes and dated, small-sample studies with a comprehensive review of the

---

[80] *See supra* note 38 at 62.

[81] *See* Julie Brill, Comm'r, FTC, Keynote Address at the Amsterdam Privacy Conference, Transatlantic Privacy After Schrems: Time for an Honest Conversation (Oct. 23, 2015), *available at* https://www.ftc.gov/system/files/documents/public_statements/836443/151023amsterdamprivacy1.pdf

[82] There do not appear to be market differences in the threat of FTC enforcement actions that could account for this pattern. Rather, FTC actions have focused on relatively large firms or on behavior that resulted in losses, such as data breaches.

content of modern privacy policies, I hope that this paper will help change the question being asked by regulators. The question is now the extent to which the current practices that the study uncovers are, in fact, desirable. Getting the right answer will require more investigation of the benefits and costs of what we see in the data. Consumer preferences need to be better understood, context by context, and compared to the boundaries of data collection and usage by firms—boundaries which, at this point, are often impossible to determine.

Table 1. Summary Statistics. Company characteristics include dummy variables for nonprofits and publicly traded. Product characteristics include whether the user must pay and the popularity of the website according to Alexa.com (lower numbers mean more popular). Privacy policy characteristics include a dummy for a claim of certification to a standard, the year the policy was last updated, and the length of the policy.

	Sample N contracts	All N = 261	Adult N=17	Cloud Computing N = 28	Dating N = 40	Gaming N = 20	News and Reviews N = 18	Social Networks N = 89	Special Interest Message Board N = 49
Nonprofit	N (nonmissing)	261	17	28	40	20	18	89	49
(0 - 1)	mean	0.04	0	0	0.07	0	0	0.04	0.08
	SD	0.2	0	0	0.27	0	0	0.21	0.28
Public	N	261	17	28	40	20	18	89	49
(0 - 1)	mean	0.27	0	0.61	0.23	0.3	0.44	0.26	0.16
	SD	0.45	0	0.5	0.42	0.47	0.51	0.44	0.37
Paid Service	N	261	17	28	40	20	18	89	49
(0 - 1)	mean	0.39	0.24	0.54	0.93	0.55	0.28	0.16	0.31
	SD	0.49	0.44	0.51	0.27	0.51	0.46	0.37	0.47
Alexa Rank	N	260	17	28	39	20	18	89	49
	mean	949,099	7,738	171,891	1,537,119	50,818	440,076	1,565,688	685,503
	SD	3,766,568	22,823	820,000	3,762,442	139,330	1,819,529	5,505,351	2,462,269
	min	1	50	1	4	31	29	1	31
	median	6,101	559	195	53,349	3,643	3,676	18,034	7,485
	max	34,999,650	94,753	4,352,180	18,971,368	587,265	7,730,387	34,999,650	15,303,381
Certification is Claimed	N	261	17	28	40	20	18	89	49
(0 - 1)	mean	0.3	0	0.68	0.28	0.35	0.17	0.33	0.16
	SD	0.46	0	0.48	0.45	0.49	0.38	0.47	0.37
Year Last Updated	N	219	8	28	32	17	14	80	40
	mean	2011	2011	2012	2011	2011	2011	2011	2010
	SD	1.8	1.4	0.6	2.5	0.7	1.8	1.8	1.9
	min	2004	2009	2010	2004	2010	2007	2006	2006
	median	2012	2012	2012	2012	2011	2011	2012	2012
	max	2015	2014	2013	2013	2012	2013	2015	2013
Number of Words	N	261	17	28	40	23	18	89	49
	mean	2,176	1,356	2,166	2,008	2,783	2,315	2,469	1,772
	SD	1,403	885	911	1,196	1,693	1,166	1,702	1,053
	min	9	159	442	180	529	361	241	9
	median	2,015	1,077	2,168	2,212	2,592	2,278	2,168	1,825
	max	9,368	4,262	4,031	4,462	7,749	4,631	9,368	4,209

Table 2. Correlations. Pairwise correlations; N=261 unless otherwise indicated. Popularity is the negative of the log Alexa rank. Other variables are described in Table 1.

	Nonprofit	Public	Paid Service	Popularity	Certification is Claimed	Year Last Updated	Log Number of Words
Nonprofit	1						
Public	0	1					
Paid Service	-0.08	-0.10	1				
Popularity	-0.04 N=260	0.44*** N=260	-0.10 N=260	1			
Certification is Claimed	-0.05	0.27***	0.05	0.25*** N=260	1		
Year Last Updated	0.04 N=219	0.12* N=219	-0.03 N=219	0.33*** N=218	0.36*** N=219	1	
Log Number of Words	-0.12*	0.25***	0.05	0.29*** N=260	0.36*** N=261	0.20*** N=219	1

* p<0.10, ** p<0.05, *** p<0.01

Table 3. Guidelines, Compliance, and Maximum Protection. The fraction of policies compliant with requirements. The footnote describes the definition of compliance in special cases. n/a terms are always excluded from proportions. Numbers may not add to 1.00 due to rounding.

			HEW FIPs 1973	European Safe Harbor 2000	FTC FIPs 2000	White House Privacy Bill of Rights 2012	FTC Privacy Report 2012	Maximum Privacy Protection .
Notice								
N1. Policy is accessible through a direct link from the homepage	yes no n.a.	0.88 0.11 0.01	.	.	yes[1] (0.88 comply)	.	yes[1] (0.88 comply)	yes[1] (0.88 comply)
N2. Users asked to manifest consent when signing up via clickwrap	yes no n.a.	0.19 0.8 0.01	yes[1] (0.19 comply)	.	yes[1] (0.19 comply)	.	yes[1] (0.19 comply)	yes[1] (0.19 comply)
N3. Layered or short notice is presented	yes no	0.2 0.8	.	.	.	yes (0.2 comply)	.	yes (0.2 comply)
N4. Contact data is collected and stored	yes no	0.96 0.04	must disclose (all comply)	.	must disclose (all comply)	.	.	no (0.04 comply)
N5. Computer data is collected and stored (e.g., IP address, browser type, OS)	yes no undisclosed	0.87 0.03 0.1	must disclose (0.90 comply)	.	must disclose (0.90 comply)	.	.	no (0.03 comply)
N6. Interactive data is collected and stored (e.g., browsing behavior or search history)	yes no undisclosed	0.71 0.13 0.16	must disclose (0.84 comply)	.	must disclose (0.84 comply)	.	.	no (0.13 comply)
N7. Financial information is collected and stored (e.g., account status or history, credit)	yes no n.a. undisclosed	0.47 0.23 0.03 0.27	must disclose (0.72 comply)	.	must disclose (0.72 comply)	.	must disclose (0.72 comply)	no (0.24 comply)
N8. Content is collected and stored (e.g., personal communications, stored documents, media)	yes no undisclosed	0.4 0.19 0.41	must disclose (0.59 comply)	.	must disclose (0.59 comply)	.	.	no (0.19 comply)
N9. Sensitive information is collected and stored (e.g., race, medical info, religion, sexual orientation, income, SSN)	yes no undisclosed	0.27 0.41 0.32	must disclose (0.68 comply)	.	must disclose (0.68 comply)	.	must disclose (0.68 comply)	no (0.41 comply)
N10. Geolocation information is collected and stored (not just IP address)	yes no n.a. undisclosed	0.15 0.52 0.32 0.01	must disclose (.99 comply)	.	must disclose (.99 comply)	.	must disclose (.99 comply)	no (0.76 comply)
N11. Cookies used	yes no undisclosed	0.92 0.02 0.05	.	.	must disclose (0.94 comply)	.	.	no (0.02 comply)
N12. PII used internally only for business purposes (e.g., administering transaction, communication with user, research, internal database compilation, servicing site)	yes no	0.3 0.7	yes (.30 comply)	.	yes (.30 comply)	yes (.30 comply)	yes (.30 comply)	yes (.30 comply)
N13. PII used only for stated, context-specific purposes (e.g., user would expect that this data would be shared for service to function)	yes no n.a.	0.26 0.73 0.01	yes (.26 comply)	.	yes (.26 comply)	yes (.26 comply)	yes (.26 comply)	yes (.26 comply)
N14. Profile, picture, or other information may be used in adverstising	yes no opt-in/opt-out	0.03 0.92 0.05	.	.	(no or) user's option (0.97 comply)	no (0.92 comply)	.	no (0.92 comply)
N15. Third parties may place advertisements that track user behavior	yes no undisclosed	0.62 0.14 0.24	.	.	no[2] (0.14-0.39 comply)	must disclose (0.76 comply)	no[2] (0.14-0.39 comply)	no[2] (0.14-0.39 comply)
N16. Recipients of shared or sold data are identified	yes no	0.09 0.91	yes (0.09 comply)	yes (0.09 comply)	yes (0.09 comply)	.	.	yes (0.09 comply)
N17. Words such as "affiliates" or "third parties" are defined, if used	yes no n.a.	0.07 0.79 0.15	.	yes[1] (0.08 comply)	yes[1] (0.08 comply)	yes[1] (0.08 comply)	.	yes[1] (0.08 comply)
N18. Company alerts user to material changes to the policy (simply posting new policy constitutes no notice)	yes no n.a. undisclosed	0.34 0.51 0.13 0.02	.	.	.	.	yes[3] (.39-.41 comply)	yes[3] (.39-.41 comply)
N19. User must explicitly assent to material changes	yes no n.a.	0.1 0.78 0.13	yes[1] (0.10 comply)	yes[1] (0.10 comply)	.	.	yes[1] (0.10 comply)	yes[1] (0.10 comply)
N20. Material changes are retroactive	yes no	0.06 0.94	no (0.94 comply)	.	.	.	.	no (0.94 comply)
N21. Provides notice of data procedures if company is sold or otherwise ceases to exist	yes no	0.08 0.92	yes (0.08 comply)	.	.	yes (0.08 comply)	.	yes (0.08 comply)
Sharing								
SH1. Affiliates and subsidiaries are bound by the same privacy policy	yes no n.a.	0.13 0.49 0.38	.	yes[1] (0.21 comply)	.	yes[1] (0.21 comply)	yes[1] (0.21 comply)	yes[1] (0.21 comply)

SH2. Contractors (e.g., payment process companies) are bound by the same privacy policy	yes no n.a.	0.2 0.59 0.21	.	yes[1] (0.25 comply)	.	yes[1] (0.25 comply)	yes[1] (0.25 comply)	yes[1] (0.25 comply)
SH3. Third parties are bound by the same privacy policy	yes no n.a.	0.05 0.72 0.23	.	yes[1] (0.06 comply)	.	yes[1] (0.06 comply)	yes[1] (0.06 comply)	yes[1] (0.06 comply)
SH4. Company shares PII information with affiliates	yes no	0.48 0.52	.	.	must disclose (all comply)	no (0.52 comply)	no (0.52 comply)	no (0.52 comply)
SH5. Company shares PII information with third parties	yes no	0.68 0.32	must disclose (all comply)	.	must disclose (all comply)	no (0.32 comply)	no (0.32 comply)	no (0.32 comply)
SH6. Company reports performing due diligence to ensure legitimacy of third parties that have access to data	yes no	0.03 0.97	.	.	.	yes (0.03 comply)	.	yes (0.03 comply)
SH7. Company has contract with third parties establishing how disclosed data can be used	yes no n.a.	0.08 0.73 0.18	.	yes[1] (0.10 comply)	.	yes[1] (0.10 comply)	.	yes[1] (0.10 comply)
SH8. Consent mechanism for sharing/selling PII or sensitive information (except for typical internal business purposes)	opt-in opt-out mandatory n.a.	0.26 0.07 0.36 0.31	opt-in[4] (0.38 comply)	user's option[5] (0.48 comply)	user's option[5] (0.48 comply)	user's option[5] (0.48 comply)	opt-in[4] (0.38 comply)	opt-in[4] (0.38 comply)
User Control								
UC1. User can adjust privacy settings	yes no	0.57 0.43	.	.	yes (0.57 comply)	yes (0.57 comply)	.	yes (0.57 comply)
UC2. User allowed to access and correct personal data collected	no can access and correct can access	0.29 0.66 0.06	can access and correct (0.66 comply)	can access and correct (and third parties notified) (0.66 comply)	can access and correct (0.66 comply)	can access and correct (0.66 comply)	can access (0.72 comply)	can access and correct (0.66 comply)
UC3. User can request that information be deleted or anonymized	yes no	0.56 0.43	yes (0.56 comply)	.	yes (0.56 comply)	yes (0.56 comply)	yes (0.56 comply)	yes (0.56 comply)
UC4. User given a choice of what happens to data if company is sold or otherwise ceases to exist	yes no	0.02 0.98	yes (0.02 comply)	.	yes (0.02 comply)	.	.	yes (0.02 comply)
Security								
SEC1. Guarantees data accuracy	yes no	0.02 0.98	yes (0.02 comply)	yes (0.02 comply)	yes (0.02 comply)	yes (0.02 comply)	yes (0.02 comply)	yes (0.02 comply)
SEC2. Company adopts reasonable procedures to ensure accuracy	yes no	0.33 0.67	yes (0.33 comply)	yes (0.33 comply)	yes (0.33 comply)	yes (0.33 comply)	yes (0.33 comply)	yes (0.33 comply)
SEC3. Company reserves right to disclose protected information to comply with the law or prevent a crime	yes no	0.85 0.15	.	.	.	yes (0.85 comply)	yes (0.85 comply)	no (0.15 comply)
SEC4. Company reserves right to disclose protected information to protect own rights	yes no	0.7 0.3	.	.	yes (0.7 comply)	yes (0.7 comply)	.	no (0.3 comply)
SEC5. User will be given notice of government requests for information about the user	yes no	0.03 0.97	.	.	yes (0.03 comply)	yes (0.03 comply)	.	yes (0.03 comply)
SEC6. Describes substantive privacy and security protections incorporated into operating procedures (e.g., limiting number of employees with access to data)	yes no	0.46 0.54	yes (0.46 comply)	yes (0.46 comply)	yes (0.46 comply)	yes (0.46 comply)	yes (0.46 comply)	yes (0.46 comply)
SEC7. Identifies means of technological security (e.g., encryption)	yes no	0.45 0.55	yes (0.45 comply)	yes (0.45 comply)	yes (0.45 comply)	yes (0.45 comply)	yes (0.45 comply)	yes (0.45 comply)
Data Practices								
DP1. States time limit for data retention (including when account is closed)	yes no	0.06 0.94	.	.	yes (0.06 comply)	yes (0.06 comply)	yes (0.06 comply)	yes (0.06 comply)
DP2. Policy for personal data when account is closed	destroyed or anonymized retained as if service continues retained but modified	0.09 0.1 0.19 0.63	.	.	.	must disclose (0.37 comply)	.	destroyed or anonymized (0.09 comply)
DP3. Has a procedure for safely disposing of unused data	yes no	0.01 0.99	.	yes (0.01 comply)	yes (0.01 comply)	yes (0.01 comply)	yes (0.01 comply)	yes (0.01 comply)
Enforcement								
E1. Provides contact information for privacy concerns or complaints	yes no	0.94 0.06	.	yes (0.94 comply)	yes (0.94 comply)	.	.	yes (0.94 comply)
E2. Disclaims liability for failure of security measures	yes no	0.58 0.42	.	.	.	.	no (0.42 comply)	no (0.42 comply)
E3. Provides link to FTC's Consumer Complaint Form and/or its telephone number	yes no	0.08 0.92	.	yes (0.08 comply)	.	.	.	yes (0.08 comply)
E4. Claims privacy seal, certification, or consistency with an industry oversight organization's practice	yes no	0.3 0.7	.	yes (0.30 comply)	yes (0.30 comply)	.	.	yes (0.30 comply)

Privacy By Design

PBD1. Requires periodic complicance review of structural and technological data security measures	yes	0.13	yes	yes	yes	yes	yes	yes
	no	0.87	(0.13 comply)	(0.13 comply)	(0.13 comply)	(0.13 comply)	(0.13 comply)	(0.13 comply)
PBD2. Contains self-reporting measures in case of privacy violation (to a privacy seal organization, third-party consultant)	yes	0.05	.	yes	yes	yes	.	yes
	no	0.95		(0.05 comply)	(0.05 comply)	(0.05 comply)		(0.05 comply)

Special definitions of compliance:
1. yes/(yes+no)
2. range given since undisclosed may be counted as a "no" depending on assumptions about undisclosed behavior and legal treatment
3. range given since undisclosed may be counted as a "yes" depending on assumptions about undisclosed behavior and legal treatment
4. opt-in/(opt-in+opt-out+mandatory)
5. (opt-in+opt-out)/(opt-in+opt-out+mandatory)

Table 4. Number of Contracts Compliant with Guidelines. In brackets, we examine 56 policies that claim compliance with the European Safe Harbor 2000 guidelines. Results are adjusted to account for policies for which a given term is not applicable.

	Number of Contracts (N = 261)				
Fraction of Terms Complied With	HEW FIPs 1973 (24 total terms)	European Safe Harbor 2000 (19 total terms)	FTC FIPs 2000 (35 total terms)	White House Privacy Bill of Rights 2012 (30 total terms)	FTC Privacy Report 2012 (27 total terms)
90%+	.	.	.	.	.
80-89%	1	.	.	.	1
70-79%	15	.	7	1	2
60-69%	43	1	47	1	13
50-59%	125	13 [12]	138	21	50
40-49%	65	8 [6]	61	54	80
30-39%	12	34 [15]	8	97	86
20-29%	.	80 [13]	.	69	27
10-19%	.	84 [10]	.	17	2
0-9%	.	41	.	1	.

Table 5. Compliance and Protection by Market. Panel A reports the average fraction of the 27 privacy policy terms in the FTC Privacy Report (2012) that is consistent with its guidelines. For example, in the average policy in the full sample, 39% of the 27 terms in these guidelines are satisfied (OVERALL); 45% of the ten notice-related terms in the guidelines are satisfied, etc. Panel B reports a weighted-average fraction of the 49 privacy policy terms that are maximally protective of consumer privacy. Each term is weighted according to how many benchmarks address it in Table 3 (term N1 is addressed in two benchmarks, N2 is addressed in three, etc.). For the average firm, this weighted average measure of protection is 30% across all 49 terms; 28% within the 21 notice-related terms, etc. Green (red) shading denotes that the average policy within a market is statistically significantly more (less) compliant or protective than the average policy for all markets at the 10% level.

	All N = 261	Adult N = 17	Cloud Computing N = 28	Dating N = 40	Gaming N = 20	News and Reviews N = 18	Social Networks N = 89	Special Interest Message Board N = 49
Panel A. Compliance								
OVERALL	0.39	0.53	0.45	0.38	0.34	0.38	0.38	0.34
Notice	0.45	0.68	0.44	0.46	0.41	0.5	0.44	0.41
Sharing	0.37	0.68	0.39	0.35	0.19	0.33	0.34	0.43
User Control	0.64	0.53	0.71	0.75	0.65	0.61	0.69	0.46
Security	0.42	0.51	0.63	0.39	0.41	0.41	0.41	0.33
Data Practices	0.04	0.03	0.02	0.01	0.03	0	0.07	0.03
Enforcement	0.42	0.47	0.61	0.48	0.3	0.39	0.37	0.39
Privacy by Design	0.13	0.06	0.43	0.03	0.15	0.06	0.16	0.04
Panel B. Protection								
OVERALL	0.3	0.42	0.36	0.27	0.26	0.28	0.3	0.28
Notice	0.28	0.5	0.29	0.23	0.27	0.31	0.27	0.28
Sharing	0.33	0.61	0.35	0.32	0.17	0.29	0.3	0.38
User Control	0.54	0.39	0.6	0.61	0.55	0.54	0.59	0.41
Security	0.28	0.36	0.43	0.25	0.25	0.24	0.27	0.24
Data Practices	0.04	0.03	0.02	0.02	0.03	0.01	0.07	0.03
Enforcement	0.49	0.35	0.67	0.5	0.48	0.47	0.5	0.44
Privacy by Design	0.1	0.06	0.38	0.02	0.09	0.03	0.11	0.03

Table 6. Compliance and Protection Regressions. Linear regressions where the dependent variable in the left set of columns is the fraction of terms in that category that are consistent with the FTC Privacy Report (2012) and the dependent variable in the right set of columns is the weighted average fraction of terms in that category that are maximally protective of consumer privacy. The weights are proportional to how many benchmarks address it in Table 3 (term N1 is addressed in two benchmarks, N2 is addressed in three, etc.). Popularity is the negative of the log Alexa rank (SD=3.90). Other independent variables are as summarized in Table 1. Standard errors are in parentheses. The stars in the Market Fixed Effects cells refer to the joint significance of the fixed effects.

	Dependent Variable: Compliance								*Dependent Variable: Protection*							
	OVERALL	**Notice**	**Sharing**	**User Control**	**Security**	**Data Practices**	**Enforcement**	**Privacy by Design**	**OVERALL**	**Notice**	**Sharing**	**User Control**	**Security**	**Data Practices**	**Enforcement**	**Privacy by Design**
Nonprofit	0.023	-0.009	0.204**	0.059	-0.077	0.000	-0.111	0.134	0.045	0.040	0.153*	0.034	-0.037	0.040	0.013	0.073
	(0.036)	(0.050)	(0.091)	(0.123)	(0.069)	(0.049)	(0.161)	(0.114)	(0.030)	(0.036)	(0.080)	(0.087)	(0.066)	(0.048)	(0.035)	(0.085)
Public	0.011	0.021	0.039	-0.056	-0.007	-0.002	-0.050	0.130**	0.008	-0.004	0.039	-0.020	-0.004	-0.007	-0.007	0.096**
	(0.018)	(0.024)	(0.045)	(0.060)	(0.034)	(0.024)	(0.079)	(0.056)	(0.015)	(0.018)	(0.039)	(0.043)	(0.033)	(0.023)	(0.017)	(0.042)
Paid Service	-0.019	-0.010	-0.029	-0.032	0.003	-0.014	-0.074	-0.044	-0.015	-0.014	-0.026	-0.036	0.010	-0.014	-0.011	-0.049
	(0.017)	(0.024)	(0.044)	(0.060)	(0.034)	(0.024)	(0.079)	(0.055)	(0.015)	(0.017)	(0.039)	(0.043)	(0.032)	(0.023)	(0.017)	(0.041)
Popularity	0.003	0.000	0.003	0.015*	0.002	0.001	0.010	0.009	0.005**	0.004	0.005	0.012**	0.004	0.001	0.001	0.010*
	(0.002)	(0.003)	(0.006)	(0.008)	(0.004)	(0.003)	(0.010)	(0.007)	(0.002)	(0.002)	(0.005)	(0.006)	(0.004)	(0.003)	(0.002)	(0.005)
Certification is Claimed	0.062***	0.046*	0.063	0.087	0.046	-0.020	0.232***	0.224***	0.074***	0.044**	0.057	0.067	0.084**	-0.015	0.355***	0.174***
	(0.018)	(0.025)	(0.045)	(0.061)	(0.034)	(0.025)	(0.080)	(0.056)	(0.015)	(0.018)	(0.040)	(0.043)	(0.033)	(0.024)	(0.018)	(0.042)
Year Last Updated	-0.010**	-0.000	-0.011	-0.029*	-0.023***	0.010*	-0.013	-0.020	-0.009**	-0.003	-0.012	-0.027**	-0.018**	0.010*	0.000	-0.015
	(0.004)	(0.006)	(0.011)	(0.015)	(0.009)	(0.006)	(0.020)	(0.014)	(0.004)	(0.004)	(0.010)	(0.011)	(0.008)	(0.006)	(0.004)	(0.011)
Log Number of Words	0.030**	0.036**	-0.145***	0.185***	0.149***	0.034**	-0.234***	-0.003	0.004	-0.011	-0.117***	0.165***	0.055**	0.030*	0.015	-0.003
	(0.012)	(0.017)	(0.032)	(0.043)	(0.024)	(0.017)	(0.056)	(0.039)	(0.011)	(0.012)	(0.028)	(0.030)	(0.023)	(0.017)	(0.012)	(0.029)
Market Fixed Effects	Yes***	Yes***	Yes**	Yes***	Yes***	Yes	Yes	Yes	Yes***	Yes***	Yes***	Yes***	Yes*	Yes	Yes	Yes**
N	218	218	218	218	218	218	218	218	218	218	218	218	218	218	218	218
Adj R-squared	0.21	0.10	0.17	0.18	0.25	0.02	0.09	0.20	0.26	0.20	0.16	0.23	0.11	0.02	0.75	0.25

* p<0.10, ** p<0.05, *** p<0.01

Table 7. FTC 2000 Versus 2012 Compliance Regressions. Linear regressions where the dependent variable is the fraction of terms in that category that are consistent with terms in the FTC FIPS (2000) and FTC Privacy Report (2012) guidelines. The first column dependent variable is the fraction of compliance with the fifteen terms unique to the 2000 guidelines (i.e., dropped from the 2012 guidelines). The second column dependent variable is the fraction of compliance with the seven terms new in the 2012 guidelines. The third column dependent variable is the fraction of compliance with sixteen terms found in both guidelines. Four additional terms included in both guidelines are dropped from consideration because the definition of compliance changed. All control variables and market fixed effects in Table 6 are included. Updated After 2012 is a dummy variable for policies last updated after the issuance of the 2012 guidelines. The sample excludes policies last updated in the year 2012. Standard errors are in parentheses.

	Compliance With FTC Guidelines		
	Terms Unique to 2000 Guidelines	**Terms Unique to 2012 Guidelines**	**Terms in Both Guidelines**
Updated After 2012	-0.002 (0.013)	-0.053 (0.047)	-0.043 (0.025)
Controls and Market F.E.s	Yes	Yes	Yes
N	125	125	125
Adj R-squared	.41	-.06	.20

* p<0.10, ** p<0.05, *** p<0.01

Figure 1. Compliance With FTC 2012 Guidelines