Self-Regulation and Competition in Privacy Policies

Florencia Marotta-Wurgler

ABSTRACT

I investigate alternative explanations for the content of privacy policies. Under one model of self-regulation, firms signal their privacy protections to consumers by highlighting compliance with third-party guidelines. However, in a sample of 249 policies, only 27 percent claim compliance with a specific guideline, and the policies that do claim compliance with at least one guideline are generally inconsistent with its requirements. Alternatively, under a market-based mechanism, firms incorporate consumers' preferences directly. Consistent with this influence, there are several intuitive differences in terms across markets. Adult sites—none of which claim certification—are much more likely to give concise and clear notice of privacy practices and limit data sharing with third parties, while cloud-computing sites are particularly likely to follow stringent data security standards. Overall, privacy policy content appears to be shaped at least as much by market forces as by a self-regulatory regime based on external guidelines.

1. INTRODUCTION

Individuals spend an ever-increasing amount of their time on the Internet, interacting with websites that connect them with friends or deliver their groceries. The firms that enable these experiences often collect patterns of

FLORENCIA MAROTTA-WURGLER is a Professor of Law at 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, an anonymous referee, and participants at the Privacy Law Scholars Conference, Conference on Empirical Legal Studies, Boston University Seminar in Law and Economics, Soshnick Colloquium on Law and Economics at Northwestern University Pritzker School of Law, University of Michigan Law and Economics Workshop, Coase-Sandor Institute for Law and Economics conference Contracting over Privacy, and Harvard Law School faculty seminar for helpful comments and suggestions. I would also like to thank Amanda Conley, Nicolas Heliotis, Alex Lipton, Julienne Markel, Jordan Miller, Melissa Quartner, Isaac Sasson, Luke Smith, Christopher Van Zele, and JingJing Wu for outstanding research assistance.

[*Journal of Legal Studies*, vol. 45 (June 2016)]
 0047-2530/2016/4502-0017$10.00

usage and personal information for commercial purposes, including constructing user-specific profiles to target content or advertising, enhance the services they offer, or share or sell the information to third parties to do the same (Zarsky 2013).

With the exception of a few sectoral laws and some state laws, the United States lacks any systematic, substantive information privacy protection rules and regulations.[1] Instead, consumers' information has been protected by a voluntary regime articulated by the Federal Trade Commission (FTC). This regime, known as notice and choice, has been predominantly based on disclosure via privacy policies and has encouraged firms to adopt substantive fair information practices (FTC 1998). These suggested practices were most recently revised in a report to Congress (FTC 2012).

Given that firms are not mandated by current federal law even to post privacy policies (though the FTC encourages it), it is interesting that almost all commercial websites do post their privacy practices in the form of privacy policies. These contracts generally, albeit sometimes not fully, describe the firms' practices in the collection, use, sharing, and protection of consumers' information. These contracts often contain more than 2,000 words and are written with care and legal formality. Their terms have only recently been studied in detail, for the first time in a large sample, by Marotta-Wurgler (2016). The obvious next questions are Where do privacy policy terms come from? Why do some firms adopt one set of terms and others quite different terms? In this paper, I shed some light on the importance of two potential influences on privacy terms, reflecting the forces of self-regulation and market competition.

First, I investigate the efficacy of self-regulation in the form of so-called privacy seals or other third-party standards. Since the 1990s, the FTC has encouraged firms to voluntarily follow a proposed set of guidelines and concluded that self-regulation was "the least intrusive and most efficient

1. See, for example, the Gramm-Leach-Bliley Act of 1999 (Pub. L. 106-102, 113 Stat. 1338), which imposes restrictions on institutions engaged in financial activities over the use and sharing of financial information and imposes disclosure requirements, and the Fair Credit Reporting Act of 1970 (15 U.S.C. 1681), which regulates the collection, use, and dissemination of credit information. A number of state laws also protect information privacy in certain contexts. See, for example, Cal. Bus. & Prof. Code, secs. 22575–78, which require website operators to post privacy policies describing their information practices; Conn. Gen. Stat., sec. 42-471, which requires businesses that collect Social Security information in the course of their business to implement privacy protection policies; and Neb. Stat., sec. 87-302(14), which prohibits firms from making false or misleading statements in privacy policies.

means to ensure fair information practices, given the rapidly evolving nature of the Internet and computer technology" (FTC 1999, p. 6). At the time, the FTC and the Department of Commerce were concerned that direct regulation would distort information markets and hinder innovation (Clinton and Gore 1997).

Encouraged by the FTC's push for self-regulation, a number of private-sector trust certification services and online seal programs have emerged, such as TRUSTe and BBB Online. Another prominent certification standard, the US-EU Safe Harbor Agreement (SHA), is a voluntary regime outlining a set of information privacy practices that US firms could abide by to ensure compliance with more stringent European information privacy laws; it has also been used by firms as a seal to signal trustworthiness.

I study the adoption of privacy seals and claims of adherence to third-party standards, including the SHA, in the sample of 249 privacy policies from Marotta-Wurgler (2016). This sample includes all of the most important firms in seven different markets—adult, cloud computing, dating, gaming, news and reviews, social networks, and special-interest message boards—as well as numerous smaller firms. Given that the presumed mechanism of third-party standards is to announce them to users and signal that their terms are reasonable, which shortcuts the need for a detailed read, one would expect such standards to be cited, at a minimum, if they are used to shape terms. Hence, as an initial analysis of the importance of third-party standards, I simply count how often they are claimed.

In fact, only 27 percent of the policies in the sample claim compliance with at least one of 15 third-party standards. Two standards—the SHA and the TRUSTe privacy seal—dominate this subsample. The SHA seal is the most popular. Of the 66 firms citing at least one standard, 32 cite the SHA alone, and 15 cite the SHA and at least one other non-SHA seal like TRUSTe—whose only function is to certify compliance with certain standards and signal trust. Just 19 of 249 firms cite only a non-SHA seal. Clearly, references to third-party standards are not widespread enough to have much influence on privacy policy terms, at least in these seven (important) markets.

One potential bright side for the efficacy of self-regulation via third-party standards is that the usage of third-party standards or certification seals may be increasing over time. While my snapshot sample is unable to directly capture changes in policies over time, there is a strong pattern that policies updated more recently are far more likely to cite third-

party standards. Unfortunately, there is a dark side. A careful read of privacy policy terms reveals that almost all firms that claim adherence to the SHA fail to meet even half of its requirements. Given that claiming adherence to the SHA subjects the firm to clear enforcement actions for noncompliance by the FTC, and firms that claim to comply still do not comply, the promise of third-party certification standards, at least with the weak enforcement mechanisms that currently accompany them, appears to be rather limited.[2]

Second, I turn toward a search for the effects of elementary competitive forces on individual terms. This is to be distinguished from the self-regulatory third-party standards model, which suggests mostly a one-size-fits-all set of terms, and the observable effect of competition is to signal quality by adopting them wholesale. As just noted, this simply has not happened for the majority of firms.

In principle, competitive outcomes for individual privacy policy terms should reflect some balance of business needs, the nature of the product, and consumers' preferences. Of course, some information practices are driven purely by the functional nature of the product. For example, social networks and dating sites tend to collect more information because their function is to relay such information to other users. I am less interested in such differences. The most telling choices about information practices across markets are those that go beyond functionality, such as notice, or sharing with third parties for purposes beyond those that are necessary.

The empirical approach that I adopt is to look for intuitive differences in terms across markets and firms that are suggestive of competitive outcomes. There turn out to be a number of such patterns. The most notable is the broad superiority of adult websites' terms relative to those of the other six markets. Adult sites give much clearer notice of their overall privacy practices than other firms yet are an average of 40 percent shorter. Adult sites are particularly noteworthy in their limited collection of personally identifiable information and in their limited sharing of it with affiliates or third parties. This would seem a straightforward outcome of competition since, given the nature of the services, an adult site

2. While the Federal Trade Commission (FTC) has brought three actions against certification providers for failure to monitor adherent firms, it has not brought any actions against individual firms for failing to comply with certifications standards. See, for example, *Federal Trade Commission v. ControlScan, Inc.,* No. 1:10-cv-00532 (March 8, 2010); *In the Matter of Facebook, Inc.,* FTC File No. 092 2184 (July 27, 2012). See also *In the Matter of True Ultimate Standards Everywhere, Inc.* (*TRUSTe*), FTC File No. 1323219 (November 17, 2014).

that routinely sold or shared its users' personally identifiable information would be at a competitive disadvantage relative to an otherwise similar site that kept users' data private. The same market force is not nearly as strong among users of gaming or news and review sites, on the other hand, which collect information and relate to activities that consumers might care less to keep private.

Another prominent and intuitive difference across markets is the stronger data security measures offered by cloud-computing sites. Users entrusting numerous personal and professional files to the cloud will naturally do so only if they are confident that their files will not be lost or stolen. Again, it is not surprising that cloud-computing firms would compete on this dimension.

To summarize, this paper makes some initial progress in explaining the wide variation across firms in the terms offered by privacy policies. On one hand, self-regulation via privacy seals and third-party standards, including the SHA, has had limited success in that most firms do not reference such standards and even those that do claim compliance with the SHA do not comply. On the other hand, elementary competitive forces are having a more detectable impact. There remains a great deal of unexplained variation in privacy terms for future work to address, however.

Section 2 reviews privacy seals and the SHA as well as the previous literature on them, introduces the sample, and reports empirical findings regarding third-party standards. Section 3 studies market and firm differences in privacy policy terms. Section 4 concludes.

2. SELF-REGULATION USING THIRD-PARTY STANDARDS

2.1. Privacy Seals

Encouraged by the FTC's push for self-regulation, a number of private-sector trust certification services and online seal programs have emerged, such as TRUSTe and BBB Online, the two largest.[3] Others include Net-

3. TRUSTe is a nonprofit created in 1997. It requires its adherents to abide by certain rules regarding notice, choice, access, and security of information. TRUSTe claims to monitor compliance and offers a dispute resolution procedure. The license agreement also requires licensees to submit to monitoring and oversight by TRUSTe as well as a complaint resolution procedure (see TRUSTe, Privacy Dispute Resolution Program [https://www.truste.com/consumer-resources/dispute-resolution/]). BBBOnLine, created in 1999, is a subsidiary of the Council of Better Business Bureaus. To obtain the organization's privacy seal, subscribers must post privacy policies in compliance with the privacy prin-

work Solutions and GeoTrust. There are also a number of security-related seals, including Norton Secured Seal by Symantec and McAfee Secure Seal, that deal exclusively with security and are used mostly by sites that involve payments.

Certification organizations require member firms to adhere to specific privacy and security codes of conduct and agree to be monitored by such organizations to ensure compliance. In exchange, compliant firms can display a seal on their websites or their privacy policies to signal a trustworthy commitment to privacy protection.[4] To register with any seal program, firms must pay an annual fee that is often based on revenue.

Regulators originally saw promise in these programs. They were hopeful that, if widely adopted, certifications would provide a clear signal or shortcut assurances of quality that consumers would find useful. And if firms in the programs were effectively monitored, these organizations could ensure good privacy practices and provide enforcement mechanisms to protect privacy.[5] But what would push certification organizations to adopt substantive information practices and entice firms to adhere to them?

Swire (1997) and others maintain that self-regulation could take off if firms feared direct regulation. Competition could also result in well-functioning certification systems if privacy was salient to consumers. Yet, absent these forces, a lack of mandatory rules or regulatory oversight could also result in systems lacking transparency and accountability, and the rules created would likely be only minimally protective as a way to stave off regulation (Hoofnagle 2016). According to Kang (1998) and Schwartz (2000), firms claiming to adhere would similarly lack the incentive to do so, which would result in low take-up rates.

Research into the effectiveness of online seal certifications and take-up

---

ciples outlined by the organization, participate in a consumer dispute resolution mechanism, and agree to be monitored by the organization (see Council of Better Business Bureaus, BBB Dispute Handling and Resolution [http://www.bbb.org/bbb-dispute-handling-and-resolution/]).

4. Other examples of voluntary self-regulatory guidelines are the Network Advertising Initiative Principles, which are based on those of the now defunct Online Privacy Alliance.

5. The FTC (1999, p. 6) expressed cautious optimism for seals: "In addition, several significant and promising self-regulatory programs, including privacy seal programs, are underway. . . . In addition, the seal programs discussed below currently encompass only a handful of all Web sites. Thus, it is too early to judge how effective these programs will ultimately be in serving as enforcement mechanisms to protect consumers' online privacy."

rates is mixed. Some studies conclude that they have been successful. Culnan (1999, 2000) and Miyazaki and Fernandez (2000) find gradual improvement in certification firms' compliance with FTC privacy guidelines. Bamberger and Mulligan (2010, p. 263) interviewed chief privacy officers at top technology firms and report that, in stark contrast to the lack of regulatory initiative in this area, the private sector has fully embraced consumer privacy protection, which has resulted in the creation of seals and certification programs that, they claim, "have been adopted widely."[6]

Yet there is a general sentiment that seals and certifications are unlikely to be a panacea. Miyazaki and Krishnamurthy (2002) find that third-party certifications have no effect on firms' information privacy practices, as revealed by their privacy policies. LaRose and Rifon (2006) find that certification services fail to adequately monitor firms' practices and award certifications to firms engaging in problematic behaviors, several of which were being sued by the FTC for privacy violations. Greenstadt and Smith (2005) question the business model of certification firms, pointing out that they need adhering firms' revenues to survive and thus employ lax privacy standards to ensure that they have subscribers. Indeed, several trust organizations have been subject to FTC and state attorney general enforcement actions for failing to keep their promises to monitor compliance with their standards, which led the FTC to express doubts about the effectiveness of such programs.[7] Edelman (2011) finds that lack of regulatory oversight of trust certifications may result in adverse selection, as sites that seek and obtain seals (at least those from TRUSTe) are significantly more likely to be untrustworthy than uncertified sites. Listokin (2015) evaluates the effect of certification on standardized privacy scores and privacy breaches over time and finds that TRUSTe certification has, if anything, a negative effect on firms' privacy protections. Despite these findings, Hoofnagle (2008) reveals that a significant number of respondents believe that firms that adhere to seals offer the highest levels of privacy protection, even if that may not be the case.

6. Bamberger and Mulligan (2010, p. 263) also state that "[s]everal self-regulatory organizations provide oversight and enforcement of voluntarily adopted privacy policies, advice, and support to businesses on privacy issues, handle consumer complaints, and monitor members' privacy commitments."

7. For example, *In the Matter of TRUSTe* (FTC No. 1323219) finds that TRUSTe failed to conduct annual recertifications of adhering companies as was promised more than 1,000 times between 2006 and 2013 and that TRUSTe misrepresented its corporate status as a nonprofit.

In my empirical analysis, I track the use of a large number of different privacy seals. The extent to which seals are used speaks to their past success and future potential. In particular, if many sites use them, then they may be influencing policy content. Still, even if many sites use them, one cannot rule out that they are window dressing, official-sounding labels for a set of terms the firms would propose anyway. Even in that case seals would have some function, however, in that consumers, once familiar with what seals stand for, would not need to take the time to review a seal user's terms. If the data show that few sites use privacy seals, however, then the conclusion is much stronger: privacy seals are having little influence.

2.2. US-EU/Swiss Safe Harbor Agreement

The SHA is a coregulatory regime that was negotiated in July 2000 by Department of Commerce and European authorities and allows US firms to comply with the more stringent requirements of European privacy laws. Each firm seeking to adhere to the SHA must register with the Department of Commerce and certify that it has complied with a number of specific requirements, including posting a privacy policy on its website with several mandatory terms and following specific privacy practices detailed in the SHA workbook.[8]

The practices that adhering firms must abide by under the SHA are mostly based on disclosure. Firms are required to include in their privacy policies details regarding the type of information collected, the purpose of its use, and the entities with whom data are shared. Firms should also provide individuals the option to disclose personal information to third parties or to use it for purposes that are different from those when the data were originally collected. In addition, firms must disclose and give individuals choices before disclosing personal information to third parties and must give individuals access to their own information and an opportunity to correct such information. Adhering firms must offer certain security protections by taking reasonable precautions to protect personal information and collect only data that are relevant for the purposes used. They must also comply with a data integrity principle by collecting data relevant for the purposes used and offer rigorous dispute resolution procedures. Finally, firms must state in their privacy policies that they adhere

8. See International Trade Administration, Safe Harbor Workbook (http://www.export.gov/safeharbor/eg_main_018238.asp).

to the SHA principles. The online appendix shows a fuller and more specific set of requirements.

The FTC plays an important role in this arrangement since it is required to monitor compliance and bring enforcement actions under section 5 of the FTC Act (15 U.S.C. 45). This allows it to police unfair and deceptive practices such as failing to comply with SHA requirements despite claiming to do so in privacy policies. Since the SHA is quite specific as to what firms must include in their privacy policies, failure to comply with this aspect of the regime is fairly straightforward. To date, the FTC has brought over 180 actions challenging unfair and deceptive information privacy practices, including 39 actions for SHA violations. Most of these have resulted in settlement agreements.[9]

Although the European Court of Justice struck down the SHA in October 2015 because it failed to adequately safeguard the information privacy of EU citizens (see *Schrems v. Data Protection Commissioner* [Case C-362/14, E.C.R. (2015)]), the FTC expressed a commitment to continue to enforce the SHA in the United States until a replacement agreement is reached. The United States and European Union are currently negotiating its replacement, called the Privacy Shield. This new program largely mimics the SHA, save for increasing protections for European citizens when data are gathered by US intelligence services and for allowing European citizens to bring actions against US firms. Importantly, the Privacy Shield envisions the same self-certification mechanism and enforcement role by the FTC.[10] Hence, understanding the effectiveness of the SHA regime is vital for current debates regarding the Privacy Shield.

Views on the effectiveness of the SHA are mixed. Kang (1998) and Schwartz (2000), among others, question its effectiveness given the FTC's limited capacity to bring enforcement actions and impose sanctions. An early and extensive study of 41 firms (Dhont et al. 2004) finds that a large number of firms claiming SHA compliance had failed to embrace all requirements and had failed to register with the Department of Commerce.

In a more recent study, however, Solove and Hartzog (2014) argue that the FTC's vigorous section 5 actions for privacy violations in the past 2 decades have led firms to comply with SHA requirements. Bamberger

9. See Federal Trade Commission, Cases and Proceedings (https://www.ftc.gov/enforcement/cases-proceedings).

10. The Privacy Shield seeks to give stronger protections to EU citizens by seeking more robust monitoring and enforcement by US agencies. The proposed regime thus includes oversight by EU regulators and the availability of dispute resolution mechanisms for EU member state citizens. See Department of Commerce (2016).

and Mulligan (2010) report that the FTC's SHA actions have driven adhering firms to hire dedicated employees to comply with the terms of the SHA and that firms' embrace of the SHA's higher standards helped them to signal their commitment to good information practices to trading partners and consumers who demand them. In this way, the SHA performs a double role: compliance with EU privacy laws and signaling trust to consumers. According to former FTC commissioner Julie Brill, the SHA has been employed by over 4,500 firms, and FTC enforcement has been "deeply effective," as section 5 gives the agency the flexibility necessary to identify problem areas that need improvement (Bracy 2015, p. 2).

I analyze the importance of the SHA for a large sample of privacy policies in two steps. First, I measure how many firms claim to comply with the SHA's requirements, as in the privacy seal analysis. I then measure the extent to which those firms actually do comply with it, on the basis of a close reading of their policies.

3. EMPIRICAL ANALYSIS

3.1. Sample

This paper uses the sample of 249 policies reviewed in Marotta-Wurgler (2016), where the sample is described in detail. In brief, all of the associated firms do business in the United States, but some have overseas operations or even headquarters. The sample policies are drawn from seven markets in which consumers often share personal or sensitive information: adult (17 sites), cloud computing (19 sites), dating (39 sites), gaming (20 sites), news and reviews (18 sites), social networks (87 sites), and special-interest message boards (49 sites). These are markets in which information sharing is typically more salient than in consumer retailer sites, in which the information practices are likely to be more salient to consumers.

Every contract was read, and the presence or absence of 49 different terms was tabulated.[11] These particular terms were chosen because they appeared in at least one information privacy guideline that has been influential, such as the Organisation for Economic Co-operation and Development's Fair Information Practice Principles, or that governs current

11. Each contract was read and graded by two law students. Any discrepancies between grading were resolved in weekly meetings with the author. Cohen's kappa measure of intergrader disparities is .88.

consumer information practices, including the latest ones introduced in 2012 by the FTC, or one of several others, including the SHA. 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. I categorize the terms as notice, sharing, user control, security, data practices, enforcement, and privacy by design.

Table 1 summarizes the characteristics of the sample firms and policies. The full-sample characteristics are as in Marotta-Wurgler (2016) and are repeated here for convenience. The sample was collected from various publicly available lists and Wikipedia, which from 2010 to 2013 had the most complete list of firms. Firms in all markets were later checked against firms listed by Alexa, a website analytics resource that tracks visits to hundreds of thousands of Internet sites and categorizes them by market to ensure that the largest, medium-sized, and smallest firms in the sample are representative of each market. The policies were collected in 2013, except for those for the adult sites, which were collected in 2015. (I explain below why the difference in collection times is of little consequence.)

Table 1 shows that in addition to involvement in different markets, the sizes and business models of sample firms vary widely. About 4 percent of sample firms are nonprofits. Only 24 percent are publicly owned; none of the adult sites are associated with public companies. The business model of about 40 percent of firms in the sample involves a paid service, such as a subscription requirement, the availability of premium access, or the ability to directly purchase goods or services. These services matter because they might affect firms' need to collect private information to make a profit and, consequently, the content of their privacy policies.

The sample sites range from household names like Facebook and Google to obscure sites like Veggiedate.com. The industry-standard ranking of website traffic is produced by Alexa. A lower number indicates a more popular site. Google, which has a privacy policy in the cloud-computing subsample and a different one in the social network subsample, has an Alexa rank of 1. Although the sample includes the largest firms in each market considered here, the median Alexa rank is 9,184. On the basis of their median Alexa rank, the adult sites are more popular.

Basic policy characteristics, distinct from their content, also vary widely. The average policy in the sample was last updated in 2011 (median, 2012). About one-sixth of firms' policies do not report the date of last update; judging from the sample collection dates, the year of last up-

Table 1. Summary Statistics

	Full Sample ($N = 249$)	Adult ($N = 17$)	Cloud Computing ($N = 19$)	Dating ($N = 39$)	Gaming ($N = 20$)	News and Reviews ($N = 18$)	Social Networks ($N = 87$)	Special-Interest Message Boards ($N = 49$)
Nonprofit (0–1) mean	.04	0	0	.05	0	0	.05	.08
Public (0–1) mean	.24	0	.42	.21	.30	.44	.26	.16
Paid service (0–1) mean	.40	.24	.79	.92	.55	.28	.16	.31
Alexa rank:								
Mean	990,400	7,738	253,313	1,538,119	50,818	440,076	1,590,688	685,503
SD	3,851,534	22,823	993,392	3,804,345	139,330	1,819,529	5,566,214	2,462,269
Min	1	50	1	4	31	29	1	31
Median	9,184	559	1,620	62,697	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
Year last updated:								
N	205	8	18	31	17	14	76	41
Mean	2011	2011	2012	2011	2011	2011	2011	2011
SD	1.8	1.4	.8	2.5	.7	1.8	1.6	2.1
Min	2004	2009	2010	2004	2010	2007	2007	2006
Median	2012	2012	2012	2012	2011	2011	2012	2011
Max	2014	2014	2013	2013	2012	2013	2013	2013

Number of words:								
Mean	2,155	1,356	2,265	2,101	2,891	2,315	2,311	1,798
SD	1,358	885	1,036	1,257	1,685	1,166	1,516	1,119
Min	9	159	442	180	529	361	241	9
Median	2,015	1,077	2,255	2,252	2,736	2,278	2,021	1,825
Max	9,128	4,262	4,031	4,462	7,749	4,631	9,128	4,909
Safe Harbor Act compliance claimed (0–1) mean	.19	0	.53	.15	.25	.11	.21	.12
Privacy seal compliance claimed (0–1) mean	.14	0	.26	.13	.20	.11	.17	.06

Note. Company characteristics include dummy variables for nonprofit and public ownership. Product characteristics include whether the user must pay and the popularity of the website according to Alexa (lower numbers indicate more popularity). Privacy policy characteristics include a dummy for a claim of compliance with the US-EU/Swiss Safe Harbor Act requirements or a privacy seal, the year the policy was last updated, and the length of the policy in words. $N = 248$ or 249 except as indicated for the year last updated.

date for almost all of these policies can be no later than 2013. The mean (median) policy in adult markets was last updated in (2011) 2012, as indicated in the summary statistics, which matches the mean and median of the full sample of firms. As a result, the difference in collection times is irrelevant, since most or all of the same adult site policies were in force 2 years before.[12]

Policies also seem long. The FTC's 2012 guidelines recommend short, streamlined, and standardized policies. The average firm's policy is 2,155 words, however, which is more than four pages of typical single-spaced text. Interestingly, the policies of adult sites are considerably shorter than those of the other six markets. This introduces the overall sample; I defer further discussion of market differences until Section 4.

3.2. Claims of Compliance

The first potential influence on privacy policies that I measure is the importance of self-regulation under the auspices of private third-party standards, as expressed by seals, certificates, and the like. The fact that there are many such standards might suggest that this is an important mechanism or influence on the content of privacy policies. A minimal test for the importance of third-party standards, and one that can be performed even without an analysis of policy content, is that the policies state that they conform to such a standard.

In this sample, which includes all of the largest firms in these markets, the mention of any third-party standard turns out to be the exception, not the norm. Table 1 indicates that only a fraction of policies in the sample claim to comply with a standard: only 14 percent mention at least one privacy seal, and only 19 percent claim to comply with the SHA's requirements, which, as noted earlier, are mostly concerned with compliance of EU privacy laws. Seven of the top 10 most popular sites (by Alexa's ranking) mention at least one standard. On the other hand, only three of the next 10 sites mention one, which is similar to the rate for the full sample. In general, the low percentage of take-up of third-party seals leads one to immediately question the breadth and effectiveness of this regime.

Table 2 illustrates the comparative importance of the third-party standards mentioned by at least one policy in the sample. The SHA and US-Swiss Safe Harbor standards are by far the most important, being cited

12. See Marotta and Svirsky (2016) for discussion of some dynamics of privacy policies.

Table 2. Certifications Claimed

Policy	*N*
US/EU or US/Swiss Safe Harbor	47
TRUSTe	20
Australian Best Practice Guidelines for Online Behavioral Advertising	4
Entertainment Software Rating Board Certification	2
International Advertising Bureau Europe EU Framework for Online Behavioral Advertising	2
Thawte certificate	2
UK Internet Advertising Bureau Good Practice Principles	2
Code Blue Security	1
Data Protection Directive 95/46/EC	1
German Laws on Privacy and Data Protection	1
GIODO (Polish chief inspector for the protection of personal data)	1
Habeas Web Seal	1
Payment Card Industry Security Standards	1
PRIVO (specializing in Children's Online Privacy Protection Act compliance)	1
UK Information Commissioner's Office	1
None claimed	183
One claimed	49
Two claimed	15
Three or more claimed	2

Note. Data are from the sample of 249 privacy policies.

more often than the 14 other standards combined. Table 2 also shows that some firms claim multiple certifications, including both SHA and privacy seals. In total, only 27 percent of firms in the sample claim one, the other, or both.

It is a bit premature to dismiss the potential of third-party standards on the basis of this simple count, however. As Figure 1 shows, policies updated more recently are much more likely to include a claim of conformity than more dated (albeit still in force) policies. For instance, 29 of the 80 policies last updated in 2012, or 36 percent, claimed conformity with at least one standard. In the most recent vintage, 24 of the 36 policies last updated in 2013, or 67 percent, claimed conformity with at least one standard.[13]

Despite appearances, one cannot definitively conclude on the basis of a snapshot cross section of privacy policies that there is a trend toward the use of third-party standards. It is mathematically possible that, for example, all of the policies that were updated in 2013 were first written in the early 2000s when the SHA was introduced, and no policies since then

13. There is only one policy dated 2014, so that vintage is not included in Figure 1.

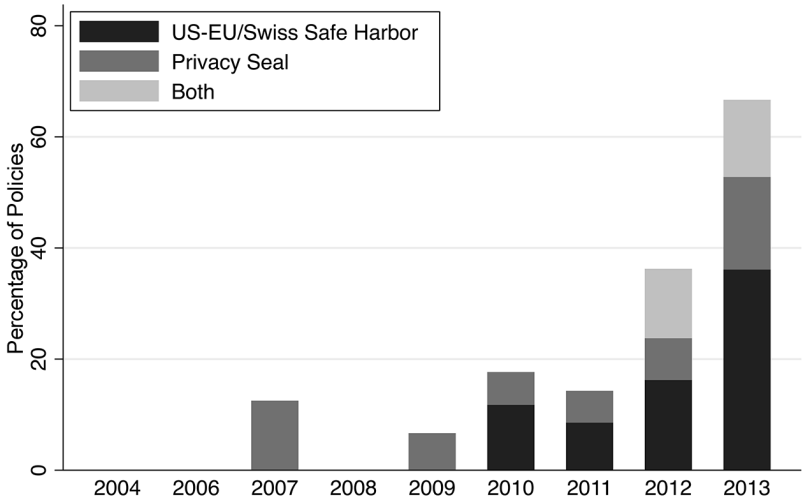

Figure 1. Claiming compliance with a third-party standard by date of last policy update

newly included the claim of SHA compliance. What can be concluded without reservation on the basis of Figure 1, however, is that firms that are actively updating their policies these days are also highly likely to be considering, and claiming, privacy seals and SHA compliance. In this sense, third-party standards have not gone out of style for firms that have updated their policies recently. On the other hand, many firms that have not updated their policies for several years do not appear to feel a pressing need to do so just to match a third-party standard.

There are two interesting cross-market differences in the popularity of citations to third-party standards. One is that cloud-computing policies are likelier than not to include mention of a third-party standard. The other is that not even one adult site cites a third-party standard.[14] Could it be that adult sites do not bother with third-party seals because their readers are more interested in the text of the policy than a claim of consistency with some unfamiliar standard? Do cloud-computing sites claim consistency with third-party standards because their users are comforted by anything that looks official, and their concern is focused on the safety

14. In unreported results, these market differences remain after controlling for year of last update, so they are not driven by, for example, an influx of newly updated cloud-computing policies.

of the data that they upload? One can speculate, but these questions are hard to answer empirically.

3.3. Claims versus Compliance

Do firms that claim compliance with a third-party guideline actually comply? That is, do their specific terms conform to that guideline? For each policy, I compute the fraction of terms that are consistent with the 19 terms in the SHA, since it is by far the most commonly claimed standard and because of its role in enabling firms to comply with EU law. The SHA also has more teeth because it requires firms that claim to adhere to it to comply or risk FTC enforcement actions, given that what constitutes a violation of it is fairly clear and, in most instances, can be observed by looking at whether the privacy policy complies with the explicit requirements. Hence, one would expect that those policies that claim SHA compliance would include terms that comply with it, perhaps not perfectly so but at least at a far higher rate than policies that do not claim SHA compliance.

Remarkably, even with a clearer threat of enforcement actions, companies that claim SHA compliance follow the SHA guidelines only to a modest degree, at least with regard to those aspects of compliance that can be ascertained by looking at a privacy policy. Figure 2 shows that firms that claim to comply with the SHA comply with only 30 percent (median, 32 percent) of its requirements. This corresponds to about six of the 19 terms that I track. Indeed, the greatest degree of term-level compliance by an SHA-claiming policy is only 58 percent, or about 11 of 19 terms. By comparison, firms that make no claim of compliance with the SHA are compliant with 18 percent (median, 16 percent) of its requirements.

To summarize, in the large sample of privacy policies analyzed here, claims of compliance with the SHA—by far the most commonly claimed standard—appear to be largely misleading window dressing. The results represent a reality check on the potential value of the Privacy Shield to ensure compliance with EU law and self-regulation via third-party standards, certificates, and seals more generally. In giving firms an easy way to falsely reassure consumers that their privacy practices are reasonable, while not including a robust mechanism for verification or enforcement, privacy seals may even be a disservice to consumers.

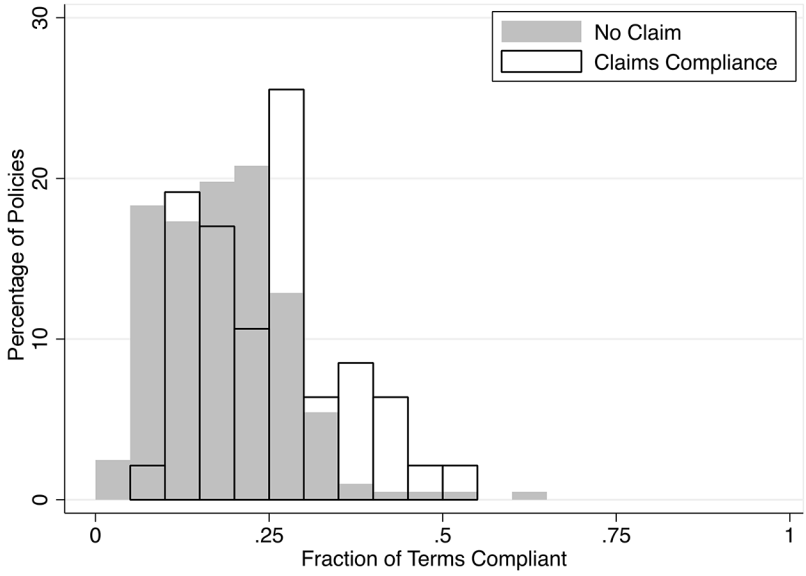

Figure 2. Firms claiming compliance with a third-party standard

4. COMPETITION OVER TERMS

The second category of influences on privacy policy content that I investigate here involves market competition over the individual terms. The idea is that firms respond to their own consumers' preferences in designing individual terms, weighing them against their particular business needs, as opposed to thinking in terms of the adoption of one-size-fits-all third-party standards. To be clear, claims of voluntary adherence to third-party standards, to the extent that such claims are not false or window dressing, can also be thought of as driven by competition; I refer to that mechanism as self-regulation to follow the common rubric.

To investigate this influence, I simply look for intuitive cross-market differences. I am aware of no natural experiments that would provide exogenous shocks to privacy-relevant competitive forces, which would be the preferred identification strategy. However, firms decide on their markets and business models prior to deciding on the fine print of privacy policies, so reverse causality is not a major concern here.

For this analysis, I benchmark privacy policies against the FTC's 2012 privacy guidelines. The FTC's 2012 guidelines involve a subset of 27 of the 49 terms for which Marotta-Wurgler (2016) collects data, and they

are listed in the online appendix. I switch to this benchmark because for non-SHA firms, which constitute the majority of the sample, the FTC's guidelines are the closest thing to a current set of formal regulatory guidelines that apply to all firms, although adherence to them remains voluntary.

4.1. Differences across Markets

Table 3 shows the average rates of compliance with the FTC's 2012 guidelines by market and term category. I measure compliance with respect to the FTC guidelines as the fraction of terms in that category that satisfies the guidelines. Compliance rates that are statistically significantly above average and below average are indicated.

The average compliance for the full sample of firms and terms in Table 3 follows Marotta-Wurgler (2016). The overall compliance measure of .38 for the full sample indicates that the average firm complies with 38 percent of the terms in the FTC's 2012 guidelines, which is roughly 10 of 27.

But differences across markets reveal some interesting and intuitive patterns. For adult sites, more than two-thirds of notice terms are compliant with the FTC's 2012 guidelines, but in no other market is the rate of compliance more than half. Despite being 800 words shorter than average, adult sites' policies provide considerably more detailed notice of many critical privacy practices. This translates to substantive protections for the consumer, as disclosures reveal little collection of consumers' personal information.

Adult sites are even more distinctive in terms of sharing practices. More than two-thirds of adult sites' sharing-related terms are compliant, while in other markets the rate of compliance varies from 17 percent to 43 percent. More important, this translates into marked differences in the level of protection. To be compliant with sharing terms tends to mean limited sharing of information, and this is the case with adult sites. If there is a market and a term category in which privacy is especially salient to consumers, this is it.

Contrast this with dating, gaming, or social networking sites. To be clear, the basic functionality of these sites 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. Associated variation in privacy policy terms does not reflect competitive pressure, just the nature of the business. Collection and sharing of this information with unknown third

Table 3. Compliance and Voluntary Protection by Market

	Full Sample ($N = 249$)	Adult ($N = 17$)	Cloud Computing ($N = 19$)	Dating ($N = 39$)	Gaming ($N = 20$)	News and Reviews ($N = 18$)	Social Networks ($N = 87$)	Special-Interest Message Boards ($N = 49$)
Overall	.38	.53[a]	.41	.38	.33[b]	.37	.37	.35[b]
Notice	.45	.68[a]	.44	.47	.41	.48	.44	.41[b]
Sharing	.36	.68[a]	.35	.34	.17[b]	.31	.34	.43
User control	.63	.53	.66	.74[a]	.63	.58	.68	.48[b]
Security	.41	.51[a]	.57[a]	.39	.42	.40	.39	.36[b]
Data practices	.04	.03	.03	.03	.03	0	.06	.03
Enforcement	.40	.47	.42	.48	.25	.39	.37	.41
Privacy by design	.10	.06	.21	.03	.15	0	.16[a]	.04[b]

Note. Values are the fraction of privacy policy terms consistent with Federal Trade Commission's 2012 guidelines. For example, the guidelines specify requirements for 10 notice-related terms, and the average privacy policy complies with somewhat more than four of these (.45).

[a] Average policy is statistically significantly more compliant than that for all markets (10% level, two-sided test).

[b] Average policy is statistically significantly less compliant than that for all markets (10% level, two-sided test).

parties, however, are not required for the site's functionality. This sort of sharing would not be as salient to consumers as it is for adult sites (Stutzman, Gross, and Acquisti 2012), which thus weakens any market demands for increased protections. In the case of gaming, liberal sharing practices drive down its overall performance to a statistically significantly below average level.

Adult sites appear to take a less pro-privacy stance in terms of users' control of privacy settings or other aspects of their information. This is somewhat misleading, however, because users generally collect and share few data in the first place, so there are few data to control. Message boards tend to limit the ability of consumers to delete or anonymize their information, which contributes to their significantly lower rate of compliance and protections here. This may be to deter anonymous "trolls" who diminish the reputation of a message board. The higher degree of compliance by dating sites naturally reflects the need to ensure current and correct information about personal characteristics, relationship status, and the like. (This is another example of variation that is simply due to the function of the sites as opposed to market forces.)

Another intuitive result involves cloud computing. Cloud-computing policies comply with the FTC's data security and security-related guidelines to a greater extent than policies for other markets. This makes sense as a basic competitive outcome given that the security of data storage is a fundamental expectation of cloud-computing users. In particular, cloud-computing sites are more explicit about their security measures. They also seek to court business users, who have valuable information at stake and might be carefully shopping around for firms with good security practices. As mentioned earlier, cloud-computing firms are also much more likely to claim compliance with third-party certifications (although in the case of SHA compliance claims these are often empty promises). The relevant privacy-by-design guideline included in the FTC's 2012 guidelines, to require periodic reviews of data security measures, is widely ignored, but in relative terms the highest rate of compliance is found in cloud-computing firms.

Figure 3 shows the distribution of overall compliance with the FTC's 2012 guidelines for each market. The graphs show the distribution for each market overlaid on the distribution for the other markets in the sample (solid bars). The main pattern that jumps out is the overall compliance rates of adult sites, which considerably exceed the average compliance of other sites. This pattern is not driven by outliers. All of the

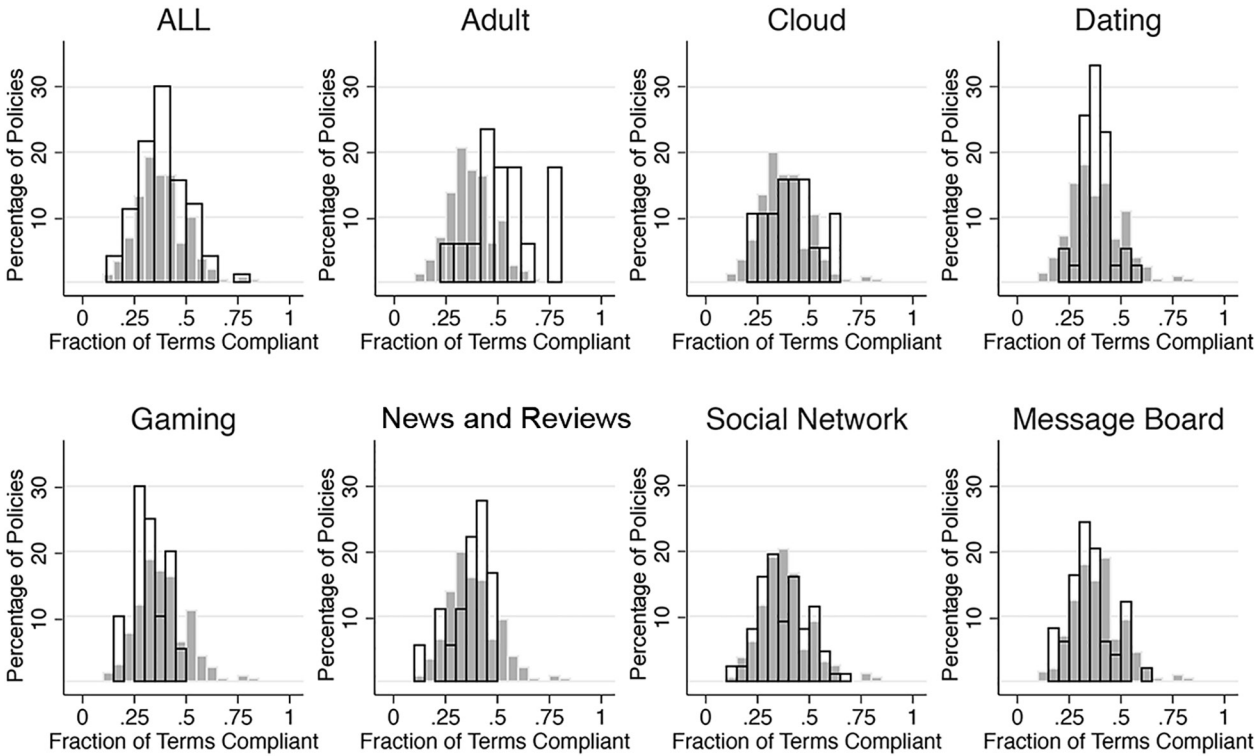

Figure 3. Compliance with Federal Trade Commission's 2012 guidelines by market

highest-scoring policies and none of the lowest-scoring policies are from adult sites.

To summarize, the data display a number of intuitive patterns consistent with competitive forces incorporating consumer preferences. Perhaps the most notable are the privacy policies of adult sites. They are the best in the sample on multiple dimensions and overall. They give particularly clear notice of their privacy practices and share far less data than any other market, despite being considerably shorter, on average, than the policies on all other sites and despite a complete absence of claims of adherence to third-party standards. Cloud-computing sites score very high for security provisions. These are patterns that competition would be expected to generate.

4.2. Differences within Markets

The final analysis looks for differences within markets. The patterns here speak less directly to competitive outcomes and resemble those that are interpreted in more detail in Marotta-Wurgler (2016); their purpose here is more to confirm the robustness of previous findings. Table 4 shows regressions of rates of compliance with the FTC's 2012 guidelines on a number of firm and site characteristics. To isolate differences within markets, the regressions include market fixed effects.

One robustness-related result is that the pattern seen in Figure 2—that claims of compliance with a privacy seal or the SHA are in fact associated with only moderate increases in overall privacy protections—are also apparent when using the FTC's 2012 guidelines as a benchmark, as done here, and when controlling for market and other firm and contract characteristics. Another result is that the market differences remain after controlling for firm and contract characteristics. As in Table 3, there remain significant differences across markets in both overall compliance and compliance in several categories of terms. For parsimony I do not report the market fixed effects, but the differences remain roughly as large as those in Table 3 and can account for at least half of the adjusted R^2-value in the overall compliance measures.

5. CONCLUSIONS

Privacy policy terms govern the relationship between consumers and firms regarding information privacy. They are of both academic and regulatory interest, but there has been little systematic study of their origins.

Table 4. Compliance

	Overall	Notice	Sharing	User Control	Security	Data Practices	Enforcement	Privacy by Design
Nonprofit	.040	.017	.172^{+}	.027	−.051	.078	.045	.156
	(.036)	(.053)	(.096)	(.132)	(.074)	(.053)	(.171)	(.117)
Public	.004	.027	−.001	−.080	−.005	.036	−.038	.051
	(.018)	(.026)	(.048)	(.065)	(.037)	(.026)	(.085)	(.058)
Paid service	−.031^{+}	−.015	−.082^{+}	−.030	.002	−.003	−.129	−.021
	(.018)	(.026)	(.047)	(.064)	(.036)	(.026)	(.083)	(.057)
Log Alexa rank	−.001	.001	−.006	−.014	.006	−.001	−.003	−.006
	(.002)	(.004)	(.006)	(.009)	(.005)	(.004)	(.011)	(.008)
Year last updated	−.009*	−.001	−.007	−.027^{+}	−.018**	.004	−.016	−.011
	(.004)	(.006)	(.012)	(.016)	(.009)	(.006)	(.021)	(.014)
Log number of words	.035**	.038*	−.150**	.208**	.163**	.024	−.224**	.004
	(.013)	(.019)	(.033)	(.046)	(.026)	(.019)	(.060)	(.041)
Safe Harbor Act compliance claimed	.036^{+}	.039	−.019	.054	.024	−.006	.232*	.115^{+}
	(.019)	(.028)	(.051)	(.070)	(.039)	(.028)	(.091)	(.062)
Privacy seal compliance claimed	.044*	.018	.139*	.001	.025	−.017	.037	.136*
	(.021)	(.031)	(.056)	(.077)	(.044)	(.031)	(.101)	(.069)
Market fixed effects	Yes**	Yes**	Yes*	Yes*	Yes^{+}	Yes	Yes	Yes
Adjusted R^2	.163	.087	.188	.164	.200	−.006	.051	.071

Note. Values are the results of linear regressions in which the dependent variable is the fraction of terms in that category that are consistent with the Federal Trade Commission's 2012 guidelines. For log Alexa rank, SD = 3.92; for log number of words, SD = .82. Standard errors are in parentheses. Superscripts for the market fixed effects refer to the joint significance of those effects. $N = 204$.

$^{+}$ $p < .10$.

* $p < .05$.

** $p < .01$.

This paper uses a large, detailed data set on the content of modern Internet privacy policies to shed some light on two categories of forces that may be shaping terms, self-regulation, and competition.

The investigation into self-regulation via privacy seals and claims of compliance with the SHA shows that policies that are fresher, in the sense of more recently updated, are much more likely to claim compliance with a third-party standard. However, because policies are not updated very frequently, only about one-quarter of privacy policies in force today—and no policies of adult sites—contain such a claim. As a result, third-party standards could be having at best limited influence on privacy policy terms.

Furthermore, policies that claim to comply with the SHA's requirements—the most commonly claimed standard and the model for the new Privacy Shield to govern data transfers between the United States and the European Union—are, on a close reading of the text of the policies, usually far from compliant. This is despite the fact that falsely claiming SHA compliance invites FTC enforcement actions. This finding further diminishes any suggestion that third-party or coregulatory standards (in the case of the SHA) are influencing policy terms, and it should concern academics and regulatory bodies that have embraced this model. It appears that firms can and often do put official-looking badges on their websites or privacy policies that have the potential to falsely reassure consumers into thinking that their privacy practices conform to a vetted external standard.

The investigation into competition over individual terms as opposed to one-size-fits-all third-party standards yields several interesting findings. Consistent with the incorporation of users' preferences, adult sites give very clear notice and have very restrictive data-sharing policies. Gaming sites are the worst. Cloud-computing sites and adult sites claim to have the best data security practices. Such intuitive cross-market differences are consistent with firms competing to meet consumers' preferences in their privacy policy terms.

REFERENCES

Bamberger, Kenneth A., and Deirdre K. Mulligan. 2010. Privacy on the Books and on the Ground. *Stanford Law Review* 63:247–315.

Bracy, Jedidiah. 2015. How Julie Brill Is Cultivating a Defense of the U.S. Privacy Framework. *Privacy Perspectives,* February 24. https://www.ftc.gov/system

/files/documents/public_statements/630801/150224juliebrillcultivatingprivacy.pdf.

Clinton, William J., and Albert Gore, Jr. 1997. *A Framework for Global Economic Commerce*. Washington, DC: White House.

Culnan, Mary J., and Pamela K. Armstrong. 1999. Information Privacy, Procedural Fairness and Impersonal Trust: An Empirical Investigation. *Organization Science* 10:104–15.

———. 2000. Protecting Privacy Online: Is Self-Regulation Working? *Journal of Public Policy and Marketing* 19:20–26.

Department of Commerce. 2016. Fact Sheet: Overview of the EU-U.S. Privacy Shield Framework for Interested Participants. July 12. https://www.commerce.gov/sites/commerce.gov/files/media/files/2016/fact_sheet-_eu-us_privacy_shield_7-16_sc_cmts.pdf.

Dhont, Jan, María Verónica Pérez Asinari, Yves Poullet, Joel R. Reidenberg, and Lee A. Bygrave. 2004. *Safe Harbour Decision Implementation Study*. Brussels: European Union Internal Market Directorate-General. http://ec.europa.eu/justice/policies/privacy/docs/studies/safe-harbour-2004_en.pdf.

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

FTC (Federal Trade Commission). 1998. *Privacy Online: A Report to Congress*. Washington, DC: Federal Trade Commission.

———. 1999. *Self-Regulation and Privacy Online: A Report to Congress*. Washington, DC: Federal Trade Commission.

———. 2012. *Protecting Consumer Information in an Era of Rapid Change: A Report to Congress*. Washington, DC: Federal Trade Commission.

Greenstadt, Rachel, and Michael Smith. 2005. Protecting Personal Information: Obstacles and Directions. Paper presented at the Fourth Workshop on the Economics of Information Security. Cambridge, MA, June 2–3.

Hoofnagle, Chris Jay, and Jennifer King. 2008. Research Report: What Californians Understand about Privacy Offline. Working paper. University of California, School of Information, Berkeley.

———. 2016. *Federal Trade Commission Privacy Law and Policy*. New York: Cambridge University Press.

Kang, Jerry. 1998. Information Privacy in Cyberspace Transactions. *Stanford Law Review* 50:1193–1294.

LaRose, Robert, and Nora Rifon. 2006. Your Privacy Is Assured—of Being Disturbed: Websites with and without Privacy Seals. *New Media and Society* 8:1009–29.

Listokin, Siona. 2015. Industry Self-Regulation of Consumer Data Privacy and Security. *John Marshall Journal of Information Technology and Privacy Law* 32:15–32.

Marotta, Florencia, and Daniel Svirsky. 2016. Do FTC Enforcement Actions Matter? Compliance before and after U.S.-E.U. Safe Harbor Agreement Actions.

Unpublished manuscript. New York University, School of Law, New York.
Marotta-Wurgler, Florencia. 2016. Understanding Privacy Policies. Unpublished manuscript. New York University School of Law, New York.
Miyazaki, Anthony D., and Ana Fernandez. 2000. Internet Privacy and Security: An Examination of Online Retailers. *Journal of Public Policy and Marketing* 19:54–61.
Miyazaki, Anthony D., and Sandeep Krishnamurthy. 2002. Internet Seals of Approval: Effects on Online Privacy Policies and Consumer Perceptions. *Journal of Consumer Affairs* 36:28–49.
Schwartz, Paul M. 2000. Internet Privacy and the State. *University of Connecticut Law Review* 32:815–59.
Solove, Daniel J., and Woodrow Hartzog. 2014. The FTC and the New Common Law of Privacy. *Columbia Law Review* 114:583–676.
Stutzman, Fred, Ralph Gross, and Alessandro Acquisti. 2012. Silent Listeners: The Evolution of Privacy and Disclosure on Facebook. *Journal of Privacy and Confidentiality* 4(2):7–41.
Swire, Peter P. 1997. Markets, Self-Regulation, and Government Enforcement in the Protection of Personal Information. Chap. 1, sec. A, in *Privacy and Self-Regulation in the Information Age*. Washington, DC: National Telecommunications and Information Administration Office of the Chief Counsel. http://www.ntia.doc.gov/reports/privacy/selfreg1.htm.
Zarsky, Tal. 2013. Transparent Predictions. *University of Illinois Law Review*, pp. 1503–69.