The John Marshall Journal of Information Technology & Privacy Law

Volume 32 | Issue 1 Article 2

Fall 2015

Industry Self-Regulation of Consumer Data Privacy and Security, 32 J. Marshall J. Info. Tech. & Privacy L. 15 (2015)

Siona Listokin

INDUSTRY SELF-REGULATION OF CONSUMER DATA PRIVACY AND SECURITY

SIONA LISTOKIN*

ABSTRACT

Industry self-regulation of consumer data privacy and security has been proposed as a flexible alternative and compliment to traditional government regulation. This study analyzes whether different types of existing industry-led standards improve online privacy and security. This paper examines which types of firms join voluntary standards and whether there is a difference in outcomes between trade association memberships (like the Digital Advertising Alliance) and certification programs (like TRUSTe). Results suggest that more trafficked websites are more likely to adopt standards, and that trade association membership does not have an effect on privacy and security performance. This article highlights the need for a valid privacy metric for robust empirical study of data privacy and security.

INTRODUCTION

As the creation and collection of consumer digital information continues its astonishing growth, consumer data privacy and security stand out as pressing areas of opportunity and concern for the online marketplace. In 2009 and 2012, the Federal Trade Commission (FTC) issued guidelines and best practices for self-regulation of consumer privacy and noted several industry-led initiatives covering the use of digital information that seek to foster innovation while protecting data privacy and security.[1] Federal legislation proposals, such as the White

* Associate Professor, School of Policy, Government and International Affairs, George Mason University, 3351 Fairfax Drive, Arlington, VA 22201, 703-993-9756, slistoki@gmu.edu.

I am grateful for support from the George Mason Law School Law and Economics Center, as well as a research grant from Google, Inc. All ideas and errors are my own.

1. *See generally* FED. TRADE COMM'N, FTC STAFF REPORT: SELF-REGULATORY

House's Consumer Privacy Bill of Rights Act of 2015, relies heavily on "enforceable codes of conduct developed by diverse stakeholders," that would surely include industry organizations.[2]

There are a number of advantages to self-regulation in fast-changing industries like "e-commerce" (broadly defined) that collect and use consumer data. Information technology is fast-changing by nature and regulatory responses may not keep pace with the industry. When properly managed, self-regulation through trade associations and certification programs can adapt more quickly and appropriately to innovations than government regulation. Moreover, self-regulation can provide a market solution to information asymmetries between firms and consumers by differentiating companies' data privacy and security performance. Firms can use industry standards to increase consumer trust without stifling the creation of new products or shifting away from a free, advertising revenue-based business model. At the same time, self-regulation typically relies on self-policing for enforcement, creating conditions for adverse selection of firms that agree to comply with standards and moral hazard for firms once they are certified. Indeed, in May 2014, the FTC called for additional guidelines for "data brokers" to protect consumer privacy, in part, because monitoring and enforcement by self-regulating organizations have not sufficiently addressed regulators' concerns.[3]

This study analyzes whether different types of existing industry-led standards for consumer data privacy and security effect online privacy. This paper examines which types of firms adopt voluntary standards, and whether there is a difference in outcomes between trade association memberships (like the Digital Advertising Alliance) and certification programs (like TRUSTe). Results suggest that more trafficked websites are more likely to be members of standards programs, and that trade association membership does not have an effect on privacy and security performance. While there is some evidence that certification can hurt

---

PRINCIPLES FOR ONLINE BEHAVIORAL ADVERTISING: TRACKING TARGETING, AND TECHNOLOGY (Feb. 2009) [hereinafter 2009 FTC STAFF REPORT], *available at* https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-staff-report-self-regulatory-principles-online-behavioral-advertising/p085400behavadreport.pdf; *see generally* FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS (Mar. 2012), *available at* https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

2. Alex Wilhelm, *White House Drops 'Consumer Privacy Bill Of Rights Act' Draft*, TECHCRUNCH (Feb. 27, 2015), techcrunch.com/2015/02/27/white-house-drops-consumer-privacy-bill-of-rights-act-draft/.

3. *See generally* FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY, i-ix (May 2014), *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

subsequent privacy and security compared to similar non-certified websites, this result is sensitive to the privacy metric used.

This study contributes to a fairly sparse literature that evaluates the effects of membership in voluntary consumer data privacy standards. The paper also compares existing measures of online privacy and highlights the significant challenges in building a valid privacy metric. Measures of website privacy that have been used in prior studies to evaluate the market for privacy frequently diverge and/or cover different websites, and empirical studies like this one may be highly sensitive to the choice of construct. The issues examined in this paper highlight the need for more rigorous empirical study and practical knowledge about different self-regulatory options in consumer data privacy.

PRIOR LITERATURE

CONSUMER DATA PRIVACY SELF-REGULATION

There are a number of studies that examine self-regulation in consumer data privacy. Scholars have introduced frameworks for viable self-regulation in this area.[4] This literature considers self-regulation as a middle ground between a pure market model and government regulation, including industry trade associations and third-party certification. Self-regulation, or industry regulation, can create rules, play a role in enforcement, and be involved in adjudication.[5] Analytical models consider consumers with heterogeneous preferences for privacy and noisy signals from firms as to privacy risk.[6] Self-regulation (or seal of approval programs) can enhance trust in situations of isolated encounters, and may be a more efficient regime than mandatory regulation.

Empirical work in this area presents a mixed picture regarding the efficacy of industry regulation. In 2000, a study of self-regulation via website privacy notices noted that one-third of the sample websites did not post privacy policies and only 14% of those that did were compre-

---

4. *See generally* Priscilla M. Regan, LEGISLATING PRIVACY: TECHNOLOGY, SOCIAL VALUES, AND PUBLIC POLICY (1995); *see generally* Peter Swire, *Markets, Self-Regulation, and Government Enforcement in the Protection of Personal Information*, *in* PRIVACY AND SELF-REGULATION IN THE INFORMATION AGE 9 (Nat'l Telecomm. & Info. Admin. 1997); *see generally* Zhulei Tang, Yu (Jeffrey) Hu & Michael D. Smith, *Gaining Trust Through Online Privacy Protection: Self-Regulation, Mandatory Standards, or Caveat Emptor*, 24 J. MGMT. INFO. SYS. 153 (2008), *available at* http://repository.cmu.edu/cgi/viewcontent.cgi?article=1048&context=heinzworks; Robert Gellman & Pam Dixon, *Many Failures: A Brief History of Privacy Self-Regulation in the United States*, WORLD PRIVACY F. (Oct. 14, 2011), http://www.worldprivacyforum.org/wp-content/uploads/2011/10/WPFselfregulationhistory.pdf.

5. Swire, *supra* note 5 at 10.

6. Tang et al., *supra* note 5 at 7-8.

hensive.[7] Subsequent works in this area show an increase in the existence of posted policies and variability in the content and readability of these privacy notices.[8] Other studies analyze specific industry membership associations or certifications. Descriptive studies that examine the content of the rules or enforcement mechanisms are typically skeptical, though there are exceptions.[9] For example, Villafranco and Riley are generally positive about the industry rules of Network Advertising Initiative ("NAI").[10] A survey of consumers showed that only 11% of users understand the NAI opt-out initiative.[11]

I examine two different voluntary programs, TRUSTe (a third party certification) and the DAA (an industry association). There are a few notable studies that specifically test the efficacy of DAA membership or TRUSTe certification beyond the case studies and review of standards highlighted above. Komanduri et al. examine the cookies produced by the DAA and NAI opt-out mechanisms of the top 100 websites and find numerous instances of non-compliance.[12] In addition, Edelman shows that TRUSTe certified websites are more likely to be rated as untrustworthy, suggesting there is an adverse selection effect for the certification seal.[13] Miyazaki and Krishnamurthy compare 60 websites' privacy policies and find that TRUSTe certification does not improve the content of privacy policies.[14]

---

7. Mary J. Culnan, *Protecting Privacy Online: Is Self-Regulation Working?*, 19 J. PUB. POL'Y & MARKETING 20, 20 (2000), *available at* http://www.jstor.org/stable/30000484?seq=1#page_scan_tab_contents.

8. *See generally*, George R. Milne, Mary J. Culnan & Henry Greene, *A Longitudinal Assessment of Online Privacy Notice Readability*, 25 J. PUB. POL'Y & MARKETING 238 (2006) (analyzing the readability of over 300 online privacy notices).

9. *See generally* Robert Gellman, *TrustE Fails to Justify Its Role As Privacy Arbiter*, 7 PRIVACY L. & POL'Y REP. 118 (2000); *See generally* Dennis D. Hirsch, *The Law and Policy of Online Privacy: Regulation, Self-Regulation, or Co-Regulation*, 34 SEATTLE U. L. REV. 439 (2011); *See generally* Saranga Komanduri, Richard Shay, Greg Norcie, Blase Ur & Lorrie Faith Cranor, *AdChoices? Compliance with Online Behavioral Advertising Notice and Choice Requirements*, 7 ISJLP 603 (2012).

10. *See generally* John E. Villafranco & Katherine E. Riley, *So You Want to Self-Regulate? The National Advertising Division as Standard Bearer*, 27 ANTITRUST 79 (2013).

11. Aleecia M. McDonald & Lorrie Faith Cranor, *Americans' Attitudes About Internet Behavioral Advertising Practices*, *in* WPES '10: PROCEEDINGS OF THE 9TH ANNUAL ACM WORKSHOP ON PRIVACY IN THE ELECTRONIC SOCIETY at 1 (2010), *available at* http://aleecia.com/authors-drafts/tprc-behav-AV.pdf.

12. Komanduri et al., *supra* note 10 at 12.

13. Benjamin Edelman, *Adverse Selection in Online "Trust" Certifications and Search Results*, 10 ELECTRONIC COM. RES. & APPLICATIONS 17, 1 (2011), *available at* http://www.benedelman.org/publications/advsel-trust-draft.pdf.

14. *See generally* Anthony D. Miyazaki & Sandeep Krishnamurthy, *Internet Seals of Approval: Effects on Online Privacy Policies and Consumer Perceptions*, 36 J. CONSUMER AFF. 28 (2002), *available at* http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.5648&rep=rep1&type=pdf.

While there are relatively few empirical evaluations of industry self-regulation of consumer data privacy, there are multiple studies of voluntary standards in other industries.[15] Studies on self-regulation and third party certification in the financial, environmental, healthcare, food and other industries find mixed results as to the efficacy of self-regulation.[16]

Voluntary standards may serve as differentiating tools for consumers and regulators that indicate superior firm management and processes designed to protect data security or control privacy.[17] Industry association membership or certification can improve certain outcomes through informal information exchange and industry pressure. Industry associations frequently facilitate communication between members and regulators, and may host member conferences to disseminate research to firms. For example, the Digital Advertising Alliance calls for transparency about data collection, and clarifies the practical application of this principle through published cases on enhanced notice to consumers.[18] "Mimetic forces" such as social networks, the creation of guidelines and best practices can lead to compliance even in the absence of sanctions.[19]

Hypothesis 1. Members of consumer data industry associations or websites have better privacy and security than non-members.

While both industry associations and third-party certifications share many characteristics conducive to effective private governance, there are a number of key differences between these models. Trade associations typically exist as self-organized institutions designed to create collective governance structures and have incentives to establish standards and behave as a "middle ground" between traditional government regulation and the free market.[20] Thus, industry organizations

---

15. *See generally* VIRGINIA HAUFLER, A PUBLIC ROLE FOR THE PRIVATE SECTOR: INDUSTRY SELF-REGULATION IN A GLOBAL ECONOMY (2001).

16. *See generally* B. James Deaton, *A Theoretical Framework for Examining the Role of Third-Party Certifiers*, 15 FOOD CONTROL 615, 615-19 (2004); *see generally* Gilles Hilary & Clive Lennox, *The Credibility of Self-Regulation: Evidence from the Accounting Profession's Peer Review Program*, 40 J. ACCT. & ECON. 211 (2005); s*ee generally* Andrew A. King & Michael J. Lenox, *Industry Self-Regulation Without Sanctions: The Chemical Industry's Responsible Care Program*, 43 ACAD. MGMT. J. 698 (2000).

17. *See generally* Michael Spence, *Job Market Signaling*, 87 Q. J. ECON. 355 (1973); Michael W. Toffel, *Resolving Information Asymmetries in Markets: The Role of Certified Management Programs* at 2 (University of California, Berkeley, Working Paper Series, Sept. 14, 2005), *available at* http://escholarship.org/uc/item/9qh5r011.

18. *Reminder: Enhanced Notice and Choice to Consumers is a 'Shared Responsibility,'* DIGITAL ADVER. ALLIANCE BLOG (Nov. 17, 2014), http://www.digitaladvertisingalliance.org/blog.aspx?id=11-17-14.

19. King & Lenox, *supra* note 17 at 701-02.

20. *See generally* ELINOR OSTROM, GOVERNING THE COMMONS: THE EVOLUTION OF

are motivated to create standards that are sufficiently restrictive to avoid external threats like government regulation and enhance their members' profiles with policymakers.

Certification organizations do not necessarily share this common goal with their clients.[21] In order for third party certifications or seals of approval to convey a credible signal of quality, they must be independent from those seeking the certification. There is little near-term incentive for for-profit certifiers to restrict their membership.[22] In addition, certified websites are not integral to – and may not even be involved in – the creation of certification standards as they are in industry associations. It is plausible that within the same industry, trade or industry-led associations may have different impacts than certifications.

Hypothesis 2. Paid privacy certifications (seals) do not improve privacy performance.

DATA

My sample includes the top 10,000 trafficked websites in 2015 that existed in 2007 and 2010. I lose about 100 websites that did not have industry information. Summary statistics are shown in Table 1. Website traffic is an imperfect screening technique for the population of comparison sites for DAA membership or TRUSTe certification. For example, the DAA includes a number of advertising network companies that do not draw much traffic. Thus, only 1% of the websites in my sample are members of the DAA in 2015, and about 4% of the websites are TRUSTe certified despite larger member rolls.[23]

DAA AND TRUSTE

TRUSTe is a private data privacy management company, with thousands of certified firms as customers.[24] TRUSTe sells software that provides a data privacy management platform and provides assessments and certification for over 4,000 firms that meet the program requirements.[25] The certification is primarily concerned with transparency and consumer choice, and includes the requirement that firms

---

INSTITUTIONS FOR COLLECTIVE ACTION, 29, 33 (1990).

21. *See* Miyazaki & Krishnamurthy, *supra* note 15 at 22-23.

22. *See* Bob Tanner, *Independent Assessment by Third-Party Certification Bodies*, 11 FOOD CONTROL 415, 415-17 (2000) (on file with author).

23. The sample is also screened for websites that exclusively carry adult content.

24. *About TRUSTe*, TRUSTe, https://www.truste.com/about-truste/ (last visited Oct. 7, 2015).

25. *Id.*

implement commercially reasonable protections for data security.[26]

It is worth noting that one of the TRUSTe programs ("TRUSTed Data") bases its requirements, in part, on the FTC self-regulatory principles, the NAI principles and the DAA principles.[27] While the TRUSTe certification provides specific monitoring guidelines, it is not clear how it deals with non-compliance. The seal requires annual recertification, and the consumer dispute resolution service processes thousands of consumer complaints every year. In 2014, the FTC charged that TRUSTe failed to follow its guidelines for privacy seal recertifications in over 1,000 incidences between 2006 and 2013; TRUSTe eventually settled with the FTC.[28]

The Digital Advertising Alliance (DAA) is a non-profit organization made up of marketing and advertising industry associations that seeks to provide self-regulatory consumer privacy principles for internet based advertising.[29] The DAA is one of the most prominent self-regulation associations in consumer data privacy and security, but has been criticized for promoting weak data privacy programs and enforcement.[30]

The DAA was established in 2009 by several U.S. advertising associations, following the release of a Federal Trade Commission (FTC) report on "Self-Regulatory Principles for Online Behavioral Advertising."[31] It is led by the Association of National Advertisers, The American Advertising Federation, 4A's, Network Advertising Initiative, Direct Marketing Association and Interactive Advertising Bureau.[32] Originally, participating companies consisted of advertisers and third party analytics companies, but starting in 2011, DAA expanded its efforts to include social networks and non-advertising firms Enforcement is handled by the Direct Marketing Association (DMA) and the Council for Better Business Bureau (CBBB).[33]

---

26. *Certification Standards*, TRUSTe, https://www.truste.com/privacy-certification-standards/ (last visited Oct. 7, 2015).

27. *TRUSTed Data Privacy Certifications*, TRUSTe, https://www.truste.com/business-products/trusted-data/ (last visited Oct. 7, 2015).

28. *TRUSTe Settles FTC Charges it Deceived Consumers Through Its Privacy Seal Program,* FEDERAL TRADE COMMISSION, (NOV. 17, 2014), https://www.ftc.gov/news-events/press-releases/2014/11/truste-settles-ftc-charges-it-deceived-consumers-through-its.

29. Pedro G. Leon, Justin Cranshaw, Lorrie F. Cranor, Jim Graves, Monoj Hastak, Blase Ur, & Guzi Xu, *What Do Online Behavioral Advertising Disclosures Communicate to Others*, 12 (April 13, 2012), https://www.cylab.cmu.edu/files/pdfs/tech_reports/CMUCyLab12008.pdf.

30. *Id.*

31. *Id.*; 2009 FTC STAFF REPORT, *supra* note 2.

32. Leon et al., *supra* note 30 at 2.

33. *Enforcement*, DIGITAL ADVERT ALLIANCE, http://www.digitaladvertisingalliance.org/content.aspx?page=enforcement (last visited on Oct. 7, 2015).

PRIVACY METRICS

The choice of privacy metric is crucial to this study and other articles that attempt to evaluate website privacy and security. While researching website privacy and security performance, I came across a number of potential measures, including some that have been used in other research. Some of these measures are included in the summary statistics and briefly described in Table 2.[34]

I use two measures of privacy and security, TrustGauge and Privacy Rights Clearinghouse data breach records. The strengths and weaknesses of these measures are discussed below.

TrustGauge is an index that measures a 10-point "trust score" based on a number of different website factors.[35] The index is not meant to exclusively or comprehensively measure privacy; it is more a construct of trustworthiness or validity.[36] Nonetheless, the measure is based in part on website privacy policies and security. The first set of features is focused on website content and verifiability, and scores sites based on the availability of contact information, privacy statements and verified customer service response. The second group of factors concerns security, such as using secure protocols on billing pages. A final set of features measures third party certification and website traffic. I subtract the points allotted for TRUSTe certification for any certified websites (no points are added for DAA membership) but do not adjust scores for website rank. As a result, I expect larger websites to have higher scores by design, and I must control for the traffic rank. I use a dummy that equals one for websites that score above 5 (the midway point). I have two full years of TrustGauge data (2007 and 2015) for the top 10,000 ranked websites in either year and a partial sample for 2010. In 2015, only 6% of the websites in my sample are coded as untrustworthy (down from about 11% in 2007).

Privacy Rights Clearinghouse (PRC) is a California nonprofit organization that has collected information about reported data breaches since 2005.[37] Data breaches are a particular element of data privacy and security. A reported breach, or absence of a breach is not necessarily an indication of underlying security weakness. In addition, the breaches reported appear to be heavily concentrated among the most trafficked websites and firms. Most breaches included in their data involve social

---

34. Of course, many potential measures are not included in this table. Notable absences include the Web of Trust (WOT), MSCI Privacy Index, EFF's Who Has Your Back?, and Terms of Service; Didn't Read. These alternatives either don't cover enough websites, are relatively new and unknown, or are very similar to the metrics included in the paper.

35. *About The Company,* TRUSTGAUGE, http://www.trustgauge.com/about.html (last visited on Oct. 7, 2015).

36. *Id.*

37 *Chronology of Data Breaches Security Breaches 2005-Present*, PRIVACY RIGHTS CLEARINGHOUSE, (April 20, 2005), https://www.privacyrights.org/node/1398.

security numbers, account numbers and other sensitive information. The PRC collects the total number of records that were compromised in each breach, however this number is "unknown" for many of the observations. In addition, the number of breached records may not be a proper indication of the lack of data security. For example, a substantial breach in Sony Pictures Entertainment in late 2014 records only 47,000 compromised records (as compared to 101.6 million records in a 2011 breach of the PlayStation Network and Sony Online Entertainment).[38] Unfortunately, these data are not consistently reported, and I create a measure based solely on the existence of a reported breach. I restrict the PRC data to non-governmental or educational targets and create a cumulative dummy each year for having ever been breached (i.e. the 2015 dummy will be 1 if the website has had at least one breach since 2005). Less than 1% of websites report more than one breach over the time period. About 2.5% of the websites in my sample have had a reported data breach by 2015.

Considering that neither TrustGauge nor PRC were created as measures of data privacy or security, it is worth comparing these constructs to other existing indices. I test the correlations between TrustGauge and the data breach records as a check of their legitimacy. The correlation table of privacy measures in 2015 is shown in Table 3. Frankly, the results are discouraging, and potentially consequential for the broader research area that evaluates data privacy. Pairwise spearman correlations are fairly low (and sometimes negative) and generally not significant. TrustGauge and breach records do not have a particularly strong (or weak) correlation to other measures. I use these metrics because of the availability of past data. The inconsistencies between the two, and the lack of convergence among privacy measures are major limitations of empirical study in this area. While other measures may prove to be stronger with time, TrustGauge and PRC have a relatively long history and cover a range of websites and companies.

I also collect information on website traffic rank and industry, from a well-known web traffic site, ranking.com.

RESULTS

My empirical strategy is to test differences in privacy outcomes for member/certified sites over time, and to refine the validity of these comparisons to approach the counterfactual: how would these websites perform without membership or certification? I begin by comparing av-

---

38 Andrew Gertz, *2014 Data Breaches by the Numbers (and the Impact)*, SAFENET, (Febuary 12, 2015), data-protection.safenet-inc.com/2015/02/2014-data-breaches-by-the-numbers-and-the-impact/; Christopher M. Loeffler, *UK ICO fines Sony £250,000 after 2011 data breach*, LEXOLOGY, (January 31, 2013), www.lexology.com/library/detail.aspx?g=417f1536-a83e-4f82-8d11-52145f0a9bd6.

erage privacy performance of members to non-members, and then restrict the sample to the most trafficked websites in order to control for the effects of site size. Finally, I match certified websites to uncertified website and employ a difference-in-differences approach to help differentiate between selection and treatment effects of membership/certification.

Initial t-tests of privacy and security show that both DAA members and TRUSTe certified websites are more likely to be rated highly by TrustGauge, and that this difference is significant (Table 4). This result supports hypothesis 1. At the same time, these websites are also more likely to have experienced a data breach.

In order to examine whether these differences persist over time and across different types of websites, I test the difference in privacy performance in both 2007 and 2015. These results are shown graphically in Figures 1 and 2 for DAA membership and TRUSTe certification, respectively. Column 1 in both figures show the likelihood of being well-rated by TrustGauge and of having a reported data breach in both time periods along with 95% confidence intervals, for the full sample of websites in the study. For the most part, the mixed results suggested by the initial t-tests persist. Member/certified sites are significantly more likely to be rated trustworthy in 2007, though this gap shrinks by 2015. These same websites are also more likely to have been breached by 2007; this difference actually grows larger by 2015.[39]

However, further tests suggest that many of these differences are related to website traffic. Column 2 of Figures 1 and 2 show the same t-tests over time for only the top 500 trafficked websites. The results are quite different; DAA members are essentially indistinguishable from non-members in terms of TrustGauge and breaches. TRUSTe certification is similarly not a differentiator by 2015, though the 2007 gaps remain.

To better control for the effects of size and other factors, I use a difference-in-differences approach. I match member/certified websites to a control group of non-certified sites and compare the difference in privacy outcomes in 2015 and 2007 (the difference between certified and non-certified websites in 2015 minus the difference between certified and non-certified websites in 2007). The matching step is intended to provide the counterfactual for how websites would have performed if they had not been certified, by eliminating time-constant unobserved effects

39 The set of firms that are certified or members differs between 2007 and 2015. Websites both join and dropout of certification. The t-test results shown in figures 1 and 2 remain almost exactly the same if I restrict the sample to websites that are consistent members, or if only compare sites that dropout. Given the subsequent results shown in this section, this is further confirmation that certification – at any point – is correlated with other factors that impact privacy and security.

on outcomes. I use propensity score matching to create better control websites that have the same probability of being certified based on website traffic rank and industry in 2007.[40]

The difference-in-differences approach typically runs across a pre- and post-treatment period. Since both DAA membership and TRUSTe certification pre-date 2007, however, I do not have a strict "pre-treatment" observation. I therefore run two tests with different treatment groups. In model A, treated websites are those that were members/certified in both 2007 and 2015. "Consistent" membership cuts the number of certified sites for both DAA and TRUSTe to 25 and 58 websites, respectively. Model B models a pre- and post-treatment by restricting certification to those sites that were not initially certified in 2007, but were certified by 2015.

Figure 3 shows the propensity score for model A certified and non-certified sites before and after matching for both DAA and TRUSTe. Results are similar for model B. In both cases, the certified sites are quite different insofar as propensity scores than the full sample of non-certified sites; the match improves the comparison considerably.

The difference-in-differences results are shown in Table 5. The first four columns show the results for DAA membership, using TrustGauge and breaches as the outcome variables for both models A and B. The certification dummy shows the difference in TrustGauge/breach averages in 2015 minus the difference in TrustGauge/breach in 2007. For the most part, the results are not significant, suggesting that trade association membership does not impact privacy. TRUSTe has a negative impact on privacy and security, though these results are not consistent across all models. Subsequent to TRUSTe certification, websites are almost 5% less likely to be rated to trustworthy compared to the trend among non-certified sites (column 6). However, this result only holds for model B, suggesting that there is an unobserved difference between websites that were certified in 2007 to those that subsequently became certified. Certification also increases the likelihood of having been breached by 2015 compared to the trend of the control websites (columns 7 and 8). These results support hypothesis 2, that paid certifications do not improve privacy.

DISCUSSION

This study evaluates industry self-regulation and paid certification of data privacy and security. Overall, the results suggest that the populations of websites that join industry associations or get certified are different than those that do not: they are simultaneously more trust-

40 *See* Paul R. Rosenbaum & Donald B. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, 70 BIOMETRIKA 41, 1 (1983), *available at* http://biomet.oxfordjournals.org/content/70/1/41.full.pdf+html.

worthy and more likely to have had a reported data breach. However, industry association membership does not appear to improve the likelihood of being trustworthy or secure. I find some evidence that the paid certification seal TRUSTe actually hurts performance compared to similar non-certified sites, as measured by TrustGauge and PRC. These results extend a similar study that finds TRUSTe certified sites are less trustworthy.[41]

As discussed previously, the privacy and security metrics used in this study are imperfect, and are not available for a true longitudinal study. Identifying an accurate measure of data privacy and security performance is difficult, and this work demonstrates that many existing metrics diverge and cover different populations of websites. The lack of convergence in the measures may undermine implications I would draw from this current study. More importantly, the inconsistency of privacy metrics has implications for the "economics of privacy" field. Empirical tests of privacy markets frequently require independent assessments of website privacy, and the construct validity of existing metrics is suspect. This is a developing area, and new measures are still being created. At the very least, the strength of privacy measures is an area ripe for future research.

With the above caveat in mind, the current results have implications for websites, policymakers and users. Industry associations, like the DAA, are meant to generate a set of best practices and to communicate with policymakers. DAA membership does not have a discernible impact on privacy performance as measured by TrustGauge. It is also possible that the DAA does improve privacy performance or the regulatory environment, but these benefits are not exclusive to members and thus fail to differentiate members from non-members. DAA could therefore facilitate industry self-regulation without confining those results to its small membership group. Paid certifications like TRUSTe may provide process management benefits, but there is no evidence in this study that its seals improve outcomes (in fact, outcomes appear to suffer). This result is in line with the FTC's complaint that TRUSTe did not perform recertifications as promised for years.

In general, the study highlights the relative dearth of empirical evaluations of certification programs and self-regulation of data privacy. Further research is needed to create a more nuanced picture of website privacy and security performance. The need for practical assessments of privacy will grow as the public and government, specifically FTC, continue to consider data privacy as a product subject to contract, transaction and fair treatment. This study outlines a research strategy to evaluate website privacy and certifications across different contexts going forward.

---

41 Edelman, *supra* note 14.

TABLES

TABLE 1

Table 1. Summary Statistics

Variable	N	Mean	Std Dev	Min	Max
Certification/Membership					
TRUSTe (2007)	9904	0.031	0.174	0	1
TRUSTe (2015)	9904	0.037	0.188	0	1
DAA (2007)	9904	0.008	0.091	0	1
DAA (2015)	9904	0.011	0.105	0	1
Privacy Metrics (outcome var's)					
TrustGauge	9904	0.945	0.227	0	1
Ever Breached	9904	0.026	0.160	0	1
Other Privacy Metrics					
Privacy Grade Scale	974	3.723	0.690	1	5
Disconnect Expected Use	836	-0.083	0.466	-1	1
Disconnect Expected Collection	836	-0.254	0.592	-1	1
Disconnect Location Data	186	-0.667	0.566	-1	1
Disconnect Data Retention	836	-0.543	0.506	-1	1
Disconnect SSL Support	9904	-0.370	0.929	-1	1
Disconnect Heartbleed Fix	3116	0.688	0.478	-1	1
SiteAdvisor	9904	0.894	0.308	0	1

2015 data unless otherwise noted

TABLE 2

Table 2. Privacy Metrics

Name	Intent	Population	Industry Curve?
TrustGauge	TrustGauge scores websites based on their trustworthiness. Scores are based on the existence of legitmate contact information, privacy policies, secure payments, certification and traffic	1,000,000+ websites	N
Privacy Rights Clearinghouse	PRC collects information about reported data breaches that has compromised personal information. The information is pulled from a number of sources including the Open Security Foundation and the California Attorney General.	Reported data breaches since 2005	NA
Privacy Grade	Privacy Grade is created by a group of researchers at Carnegie Mellon's CHIMPS lab. It grades mobile apps based on user's privacy expectations	Over 50,00 free smartphone apps	Y
Disconnect Privacy Icons	Disconnect.me Privacy Icons shows a set of icons that summarizes important elements of privacy policies and security practices. The icons cover areas like expected use and collection of data, the use of location data, data retention policies, SSL support and Heartbleed fixes. Privacy Icons pulls information from TRUSTe's Privacy Policy Database, which is separate from their certification program	Over 5,000 websites	N
SiteAdvisor	SiteAdvisor rates websites based on their security by assessing email spam and spyware.	95% of the Web	N

TABLE 3

Table 3. Privacy metric correlations (pairwise spearman correlations 2015)

	TrustGauge	Ever Breached	Privacy Grade Scale	Privacy Grade Dummy	Disconnect Expected Use	Disconnect Expected Collection	Disconnect Location Data	Disconnect Data Retention	Disconnect SSL Support	Disconnect Heartbleed Fix	SiteAdvisor
TrustGauge											
Ever Breached	0.0688 N=9904										
Privacy Grade Scale	-0.0455 N=974	0.1553 N=974									
Privacy Grade Dummy	-0.0606 N=974	0.1282 N=974	0.9583* N=974								
Disconnect Expected Use	-0.0229 N=836	0.1 N=836	0.0293 N=269	0.02 N=269							
Disconnect Expected Collection	0.0876* N=836	-0.1162 N=836	-0.1375* N=269	-0.1438* N=269	0.2516* N=836						
Disconnect Location Data	0.0882 N=183	0.0248 N=84	-0.1296 N=84	-0.1029 N=82	0.1987* N=183	0.1452 N=183					
Disconnect Data Retention	-0.0318 N=836	-0.1806 N=836	-0.0235 N=269	-0.01 N=269	0.0289 N=836	0.0186 N=836	0.3544 N=836				
Disconnect SSL Support	-0.0287* N=9904	-0.0452 N=9904	0.0261 N=974	0.01 N=974	-0.0174 N=836	-0.0283 N=836	0.0842 N=186	0.0615 N=836			
Disconnect Heartbleed Fix	-0.206* N=3116	-0.2261* N=3116	0.0284 N=273	0.0622 N=273	-0.0267 N=279	-0.0875 N=279	-0.1073 N=64	-0.0738 N=279	0 N=3116		
SiteAdvisor	0.2098* N=9880	0.1098 N=256	-0.0864* N=966	-0.0863* N=966	0.0427 N=830	0.039 N=830	0.064 N=183	0.0898* N=830	0.019 N=9880	-0.1488* N=3107	

*p<0.05
Privacy Grade covers a larger population of smartphone applications. I matched websites to identifiable apps for the top 5,000 websites.

TABLE 4

Table 4. Ttests

DAA			
2015 Membership Status	**N**	**TrustGauge**	**Data Breach**
Member	110	0.982	0.209
Non-member	9794	0.893	0.024
Difference	9904	0.089***	0.185***
Std Error		0.029	0.015
TRUSTe			
2015 Certification Status	**N**	**TrustGauge**	**Data Breach**
Certified	309	0.984	0.081
Non-certified	9595	0.891	0.024
Difference	9904	0.093***	0.057***
Std Error		0.018	0.009

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

TABLE 5

Table 5. Difference in Difference Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Certification Type	DAA				TRUSTe			
Outcome Variable	TrustGauge	TrustGauge	Ever Breached	Ever Breached	TrustGauge	TrustGauge	Ever Breached	Ever Breached
Comparison Group	A	B	A	B	A	B	A	B
Certification (Dummy)	-0.022	-0.017	0.055	0.025	-0.303	-0.049***	0.088***	0.039***
SE	0.033	0.034	0.051	0.038	0.021	0.022	0.021	0.017
N	9704	9704	9704	9704	9704	9704	9704	9704
Certified Websites	25	93	25	93	58	268	58	268

Two different treatment groups are used. In Model A, the treated websites were members/certified in 2007 (2009) and 2015. In Model B, the treated websites are only certified in 2015.
The Certification dummy is the difference-in-difference estimate. It measures the difference between treated and control websites differences in 2015 and treated and control websites differences in 2007.
Websites are matched on industry and web traffic rank.
*p < 0.10, **p < 0.05, ***p < 0.01

FIGURES

FIGURE 1

Figure 1.
DAA Membership vs Non-Membership

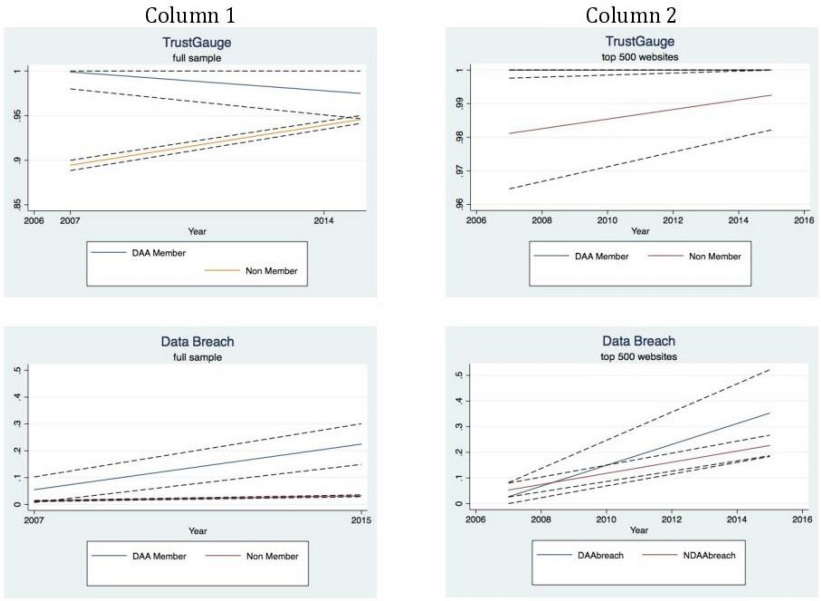


FIGURE 2

Figure 2.
TRUSTe Certification vs. Non-certification

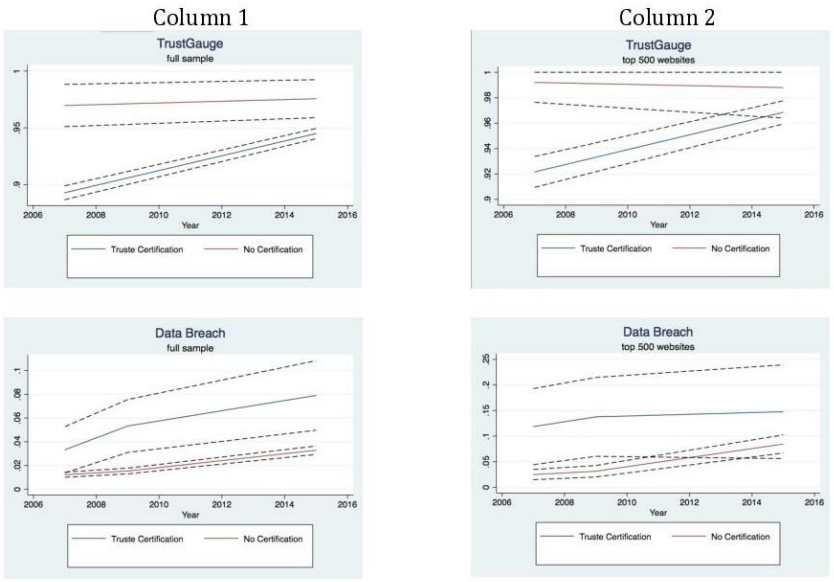
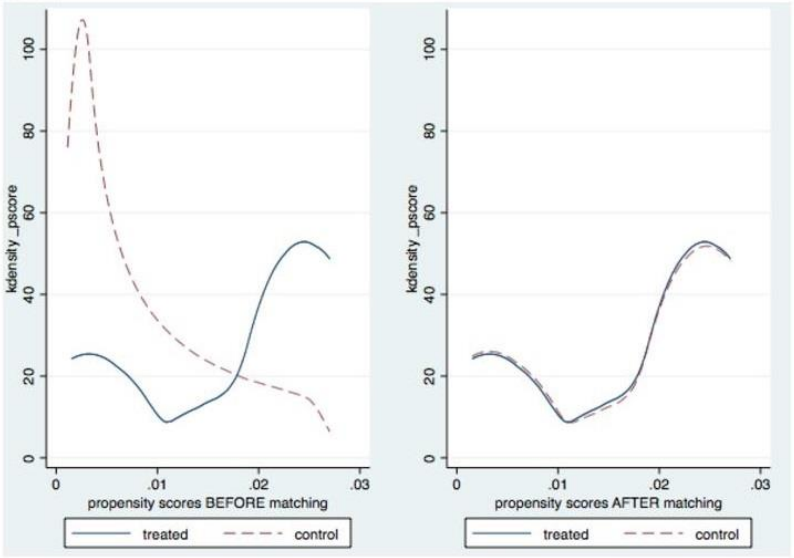


FIGURE 3

Figure 3. Propensity score before and after match

DAA



TRUSTe