Washington University Law Review

Volume 95 | Issue 1

2017

Privacy, Poverty, and Big Data: A Matrix of Vulnerabilities for Poor Americans

Mary Madden

Michele Gilman

Karen Levy

Alice Marwick

PRIVACY, POVERTY, AND BIG DATA: A MATRIX OF VULNERABILITIES FOR POOR AMERICANS

MARY MADDEN*
MICHELE GILMAN**
KAREN LEVY***
ALICE MARWICK****

ABSTRACT

This Article examines the matrix of vulnerabilities that low-income people face as a result of the collection and aggregation of big data and the application of predictive analytics. On one hand, big data systems could reverse growing economic inequality by expanding access to opportunities for low-income people. On the other hand, big data could widen economic gaps by making it possible to prey on low-income people or to exclude them from opportunities due to biases entrenched in algorithmic decision-making tools. New kinds of "networked privacy" harms, in which users are simultaneously held liable for their own behavior and the actions of those in their networks, may have particularly negative impacts on the poor. This Article reports on original empirical findings from a large, nationally-representative telephone survey with an oversample of low-income American adults, and highlights how these patterns make particular groups of low-status Internet users uniquely vulnerable to various forms of surveillance and networked privacy-related problems. In particular, a greater reliance on mobile connectivity, combined with lower usage of privacy-enhancing strategies, may

---

* Researcher, Data & Society Research Institute and Affiliate, Berkman Klein Center for Internet and Society at Harvard University. B.A., University of Florida; M.A., Georgetown University.

** Venable Professor of Law, Director of Clinical Education, Co-Director Center on Applied Feminism, University of Baltimore School of Law. B.A., Duke University; J.D., University of Michigan Law School.

*** Assistant Professor of Information Science, Cornell University; Associated Faculty, Cornell Law School. B.A., Indiana University; J.D., Indiana University Maurer School of Law; Ph.D., Princeton University.

**** Fellow, Data & Society Research Institute; Assistant Professor Department of Communication, University of North Carolina, Chapel Hill (as of 2017). B.A., Wellesley College; M.A. University of Washington; Ph.D., New York University. The authors would like to thank the participants of the 2016 Privacy Law Scholars Conference for their helpful feedback and suggestions. In particular, we are especially grateful for the feedback and guidance from Seeta Peña Gangadharan, who was the lead commentator for the paper. The authors would also like to thank the Digital Trust Foundation for funding the national survey research that supports key sections of this Article.

contribute to various privacy and security-related harms. The Article then discusses three scenarios in which big data—including data gathered from social media inputs—is being aggregated to make predictions about individual behavior: employment screening, access to higher education, and predictive policing. Analysis of the legal frameworks surrounding these case studies reveals a lack of legal protections to counter digital discrimination against low-income people. In light of these legal gaps, the Article assesses leading proposals for enhancing digital privacy through the lens of class vulnerability, including comprehensive consumer privacy legislation, digital literacy, notice and choice regimes, and due process approaches. As policymakers consider reforms, the Article urges greater attention to impacts on low-income persons and communities.

TABLE OF CONTENTS

INTRODUCTION

Low-income communities have historically been subject to a wide range of governmental monitoring and related privacy intrusions in daily life.[1] The privacy harms that poor communities and their residents suffer as a result of pervasive surveillance are especially acute in light of the resulting economic and social consequences and the low likelihood that they will be able to bear the costs associated with remedying those harms.[2] In the "big data" era, there are growing concerns that low-status Internet users who have lower levels of income or education may be further differentially impacted by certain forms of Internet-enabled data collection, surveillance, and marketing.[3] Low-status users may be both

1. *See* Michele Estrin Gilman, *The Class Differential in Privacy Law*, 77 BROOK. L. REV. 1389 (2012) [hereinafter Gilman, *Class Differential*]; *see also* Virginia Eubanks, *Want to Predict the Future of Surveillance? Ask Poor Communities*., AM. PROSPECT (Jan. 15, 2014), https://perma.cc/MQ3T-NWPB.

2. In this Article, poverty is defined as "economic deprivation," although we do not endorse any particular method of measuring poverty. *See* JOHN ICELAND, POVERTY IN AMERICA: A HANDBOOK 23 (3d ed. 2013) (defining poverty). The United States' official poverty line is an absolute measure (based on a needs standard that is constant over time), while relative measures are based on comparative disadvantage, fluctuating over time. *Id*. at 23–24. On the various methods of measuring poverty and their merits, *see generally id*. ch. 3.

3. *See* NATHAN NEWMAN, HOW BIG DATA ENABLES ECONOMIC HARM TO CONSUMERS, ESPECIALLY TO LOW-INCOME AND OTHER VULNERABLE SECTORS OF THE POPULATION (2014),

unfairly excluded from opportunities (such as access to credit) and unfairly targeted (for example, by predatory marketing strategies) based on determinations made by predictive analytics and scoring systems—growing numbers of which rely on some form of social media input.[4] These new kinds of "networked privacy" harms, in which users are simultaneously held liable for their own behavior and the actions of those in their networks, could have particularly negative impacts on the poor.[5]

In addition to the harms created by targeting or exclusion from opportunity, the poor may face magnified privacy vulnerabilities as a result of community-specific patterns around technology use and knowledge gaps about privacy- and security-protective tools.[6] Legal scholars have identified a broad group of consumers as "privacy vulnerable" when they "misunderstand the scope of data collection and falsely believe that relevant privacy rights are enshrined in privacy policies and guaranteed by law."[7] These misconceptions are common across all socioeconomic categories, but this Article suggests that these conditions may be exacerbated by poor communities' higher reliance on mobile connectivity and lower likelihood to take various privacy-protective measures online. When low-income adults rely on devices and apps that make them more vulnerable to surveillance, and they (wittingly or unwittingly) do not restrict access to the content they post online, they may be further exposed to forms of commercial data collection that can affect the way they are assessed in employment, education, and law

---

https://perma.cc/VB4Y-53SR (public comments filed in response to a Federal Trade Commission request for workshop submissions).

4. *See* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671 (2016); *see also* Jonas Lerman, *Big Data and Its Exclusions*, 66 STAN. L. REV. ONLINE 55, 56 (2013), http://www.stanfordlawreview.org/online/privacy-and-big-data/big-data-and-its-exclusions ("[B]illions of people remain on its margins because they do not routinely engage in activities that big data and advanced analytics are designed to capture.").

5. *See* danah boyd, Karen Levy & Alice Marwick, *The Networked Nature of Algorithmic Discrimination*, *in* DATA AND DISCRIMINATION: COLLECTED ESSAYS 53–57 (Seeta Peña Gangadharan, Virginia Eubanks & Solon Barocas, eds., 2014), https://perma.cc/V59G-JWDE; *see generally* Alice E. Marwick & danah boyd, *Networked Privacy: How Teenagers Negotiate Context in Social Media*, 16 NEW MEDIA & SOC'Y 1051 (2014), http://nms.sagepub.com/content/16/7/1051.

6. *See* SEETA PEÑA GANGADHARAN, NEW AMERICA FOUNDATION, JOINING THE SURVEILLANCE SOCIETY? NEW INTERNET USERS IN AN AGE OF TRACKING (2013), https://www.newamerica.org/oti/joining-the-surveillance-society/; Jennifer M. Urban & Chris Jay Hoofnagle, *The Privacy Pragmatic as Privacy Vulnerable*, BERKELEY PUBLIC LAW RESEARCH PAPER No. 2514381 (2014), http://ssrn.com/abstract=2514381 (presented to the Symposium on Usable Privacy and Security Workshop on Privacy Personas and Segmentation (PPS)).

7. Urban & Hoofnagle, *supra* note 6, at 3.

enforcement contexts.[8]

Thus, we suggest that poor people are burdened many times over by data collection and privacy intrusion. Not only are the poor subject to more surveillance than other subpopulations,[9] and at higher stakes, but in addition, poor Americans' patterns of privacy-relevant behaviors and device use open them up to greater vulnerability. We demonstrate these behavioral patterns using original empirical data from a nationally representative survey and suggest that differences like these must be considered in privacy-protective policymaking and design decisions.

This Article proceeds as follows: Part I provides a historical overview of the ways in which the poor have been subject to uniquely far-reaching surveillance across many aspects of life, and how their experiences of harm may be impacted by evolving practices in big-data-driven decision making. In using the term "poor" to signify a condition of economic deprivation, this Article recognizes that low-income people in America are a diverse and multifaceted group and that each person has his or her own individualized narrative.[10] Despite this diversity, this Article highlights a shared reality for many poor people, which is heightened vulnerability to online surveillance and associated adverse outcomes.

Part II presents new empirical findings from a nationally representative survey to highlight various technology-related behaviors and concerns that suggest low-status Internet users may be especially vulnerable to surveillance and networked privacy-related harms. By providing empirical data that demonstrates the increased vulnerability of low-income Internet users to privacy violations, we identify specific patterns of access and behavior that may help inform policy and technology design decisions.[11]

---

8. Part III of this Article discusses various case studies that highlight the scenarios in which these assessments may occur.

9. *See generally* Torin Monahan, *Regulating Belonging: Surveillance, Inequality, and the Cultural Production of Abjection*, 10 J. CULTURAL ECON. 191 (2017). "[S]urveillance manifests as a multiplicity of techniques that conjure, coalesce around, and mediate the experiences of abject subjects. Abjection signifies not only extreme need or destitution, but also a kind of social exclusion . . . ." *Id*. at 192.

10. *See* JOHN GILLIOM, OVERSEERS OF THE POOR: SURVEILLANCE, RESISTANCE, AND THE LIMITS OF PRIVACY 20–21 (2001) (describing demographic, political, physical, and regional variations among poor people); *see also* Frank Munger, *Introduction to* LABORING BELOW THE LINE: THE NEW ETHNOGRAPHY OF POVERTY, LOW-WAGE WORK, AND SURVIVAL IN THE GLOBAL ECONOMY 20 (Frank Munger ed., 2002) (asserting the importance of seeing and understanding the poor as individuals with their own narratives).

11. For more on the use of empirical data in legal scholarship, *see* Theodore Eisenberg, *Why Do Empirical Legal Scholarship?* 41 SAN DIEGO L. REV 1741 (2004); Daniel Ho & Larry Kramer,

In Part III, we show why and how this matters through a legal examination of several timely case studies that demonstrate how online activity, and the emerging use of social media data in particular, might have detrimental impacts on the poor when used in high-stakes decision-making systems. This Part explains why current legal frameworks fail to shield the poor from negative outcomes.

Finally, in Part IV, we assess major proposals for protecting personal data through the lens of class vulnerability. In other words, we evaluate how these proposals might impact poor people. We agree with other scholars that additional technical and non-technical reforms are needed to address the risks associated with the use of social media data. As policymakers consider reforms, we urge greater attention to how reforms may differentially impact low-income communities.

I. THE INTERSECTION OF PRIVACY AND POVERTY

A. Brief History of Privacy-Related Vulnerabilities and Surveillance of the Poor

Historically, the poor have had far less control over the privacy of their homes, bodies, and decisions than their more affluent counterparts.[12] In Colonial America, most towns had an "overseer of the poor" who tracked poor people and either chased them out of town or auctioned them off for free labor.[13] By the 1800s, when poorhouses became the dominant poor relief policy, the poor were warehoused in dismal quarters where they labored under the watchful eye of the "keeper."[14] Even as anti-poverty policy became more benevolent in the late 1800s, the scientific charity movement relied on "friendly visitors" to investigate the homes of the poor and exhort them to higher morals.[15] For over three centuries, surveillance in various forms has served the political purposes of "containment of alleged social contagion, evaluation of moral suitability for inclusion in public life and its benefits, and suppression of working

Introduction: The Empirical Revolution in Law, 65 STAN. L. REV. 1195 (2013).

12. *See* GILLIOM, *supra* note 10, at 23.

13. *See* WALTER TRATTNER, FROM POOR LAW TO WELFARE STATE: A HISTORY OF SOCIAL WELFARE IN AMERICA 9–10 (6th ed. 1999).

14. *See id*. at 57–61; MICHAEL B. KATZ, IN THE SHADOW OF THE POORHOUSE: A SOCIAL HISTORY OF WELFARE IN AMERICA 27–28 (10th ed. 1996).

15. *See* KATZ, *supra* note 14, at 70; TRATTNER, *supra* note 13, at 91–92.

people's resistance and collective power."[16]

The New Deal created the modern welfare state and continued this history of surveillance of the "undeserving poor" (that is, able-bodied adults who were considered capable of work).[17] In administering welfare, states devised a variety of discretionary surveillance tactics—such as midnight raids on welfare recipients' homes and moral fitness tests—designed to reduce the welfare rolls and push poor women, mostly of color, into the low-wage labor force.[18] Today, states subject single mothers who draw public assistance to drug tests, DNA testing of children, fingerprinting, extreme verification requirements, and intrusive questioning about intimate relationships.[19] Some scholars and judges have argued that higher-income Americans would object if the government treated them similarly in exchange for the valuable governmental benefits they receive, such as mortgage deductions, school loans, and child care tax credits.[20] As Justice Douglas stated in his dissent to the Supreme Court's upholding of welfare home visits, "[n]o such sums are spent policing the government subsidies granted to farmers, airlines, steamship companies, and junk mail dealers, to name but a few."[21]

The structure of the current welfare system aims to put poor women to

16. Virginia Eubanks, *Technologies of Citizenship: Surveillance and Political Learning in the Welfare System, in* SURVEILLANCE AND SECURITY: TECHNOLOGICAL POLITICS AND POWER IN EVERYDAY LIFE (Torin Monahan ed., 2006) [hereinafter Eubanks, *Technologies of Citizenship*].

17. On the New Deal division between deserving and undeserving poor, see Michele Estrin Gilman, *The Return of the Welfare Queen*, 22 AM. U. J. GENDER, SOC. POL'Y & L. 247, 257–58 (2014) and sources cited therein.

18. *See* KAARYN S. GUSTAFSON, CHEATING WELFARE: PUBLIC ASSISTANCE AND THE CRIMINALIZATION OF POVERTY 21 (2011) ("The unstated but underlying goals of the rules were to police and punish the sexuality of single mothers, to close off the indirect access to government support of able-bodied men, to winnow the welfare rolls, and to reinforce the idea that families receiving aid were entitled to no more than near-desperate living standards.").

19. *See* Kaaryn Gustafson, *Degradation Ceremonies and the Criminalization of Low-Income Women*, U.C. IRVINE L. REV. 297, 312–321 (2013) [hereinafter Gusafson, *Degradation Ceremonies*]; Khiara M. Bridges, *Privacy Rights and Public Families*, 34 HARV. J.L. & GENDER 113, 114–16 (2011) (discussing Medicaid program); Gilman, *Class Differential, supra* note 1, at 1397–1400 (discussing welfare).

20. *See* Jordan C. Budd, *A Fourth Amendment for the Poor Alone: Subconstitutional Status and the Myth of the Inviolate Home*, 85 IND. L.J. 355, 404–05 (2010); MARTHA ALBERTSON FINEMAN, THE NEUTERED MOTHER, THE SEXUAL FAMILY AND OTHER TWENTIETH CENTURY TRAGEDIES 191 (1995) ("[M]iddle-class families benefit from extensive entitlement programs, be they FHA or VA loans at below mortgage market rates or employer subsidized health and life insurance. These families receive untaxed benefits as direct subsidies.").

21. Wyman v. James, 400 U.S. 309, 332 (1970) (Douglas, J., dissenting) (quoting J. Skelly Wright, *Poverty, Minorities, and Respect for Law*, 1970 DUKE L.J. 425, 427–38 (1970)).

work.[22] Yet the low-wage workplace, where one-third of workers toil,[23] is no escape from surveillance. Employers today log computer key strokes, listen to telephone calls, review emails and Internet usage, conduct drug tests, employ mystery shoppers, watch closed-circuit television, and require psychometric and "honesty" tests as conditions of employment.[24] Employers increasingly track employee movements through GPS or radio frequency devices, which "create new streams of data about where employees are during the workday, what they are doing, how long their tasks take, and whether they comply with employment rules."[25] These sorts of tools seem to have found broad use in low-wage workplaces in particular,[26] and may be purposefully overt (rather than invisible) in order to let workers know they are being watched and to control their behavior. Other forms of surveillance are more covert; the objects of surveillance are not conscious that they are being observed. These behavioral control mechanisms can take many forms—at their most extreme, they include the use of facial recognition technology to ensure employees are smiling enough and audio recording to monitor employees' tone of voice.[27] Thus, from welfare to work and beyond, low-income people have been subject to covert and overt surveillance as tools of control.

---

22. The current welfare program is called Temporary Assistance to Needy Families (TANF) and was created by the Personal Responsibility and Work Opportunity Reconciliation Act of 1996, Pub. L. No. 104–193, 110 Stat. 2105 (1996) (codified as amended in scattered sections at 43 U.S.C. § 605(a) (2000)). The work requirements are at 42 U.S.C. § 602(a)(1)(A)(ii) (2012).

23. *See* Gilman, *supra* note 1, at 1400.

24. The Converus company offers a retina-based lie detector technology called EyeDetect, to screen potential employees. *Job Applicant and Employment Screening*, CONVERUS, http://converus.com/pre-employment-ongoing-screening/. *See generally* Kirstie Ball, *Workplace Surveillance: An Overview*, 51 LAB. HIST. 87 (2010); Alex Rosenblat, Tamara Kneese & danah boyd, *Workplace Surveillance*, DATA & SOC'Y RESEARCH INST. (2014), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2536605; Ifeoma Ajunwa, Kate Crawford & Jason Schultz, *Limitless Worker Surveillance*, 105 CAL. L. REV. (forthcoming 2017).

25. Scott R. Peppet, *Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security, and Consent*, 93 TEX. L. REV. 85, 112–13 (2010).

26. *See id*. (describing sensors on company trucks and cars and Bank of America monitoring of call-center employees); Karen E. C. Levy, *The Contexts of Control: Information, Power, and Truck-Driving Work*, 31 THE INFO. SOC'Y 160 (2015); Gilman, *Class Differential*, *supra* note 1, at 1400–03.

27. Sally Davies, *From a Frown to a Smile, the Technology That's in Your Face*, FIN. TIMES (Jan. 2, 2014), http://www.ft.com/cms/s/0/ccaac9e6-6f06-11e3-9ac9-00144feabdc0.html; Rachel Emma Silverman, *Tracking Sensors Invade the Workplace—Monitors on Workers, Furniture Offer Clues for Boosting Productivity; Switching to Bigger Lunch Tables*, WALL ST. J. (Mar. 7, 2013, 11:42 AM), http://www.wsj.com/articles/SB10001424127887324034804578344303429080678; Suzanne McGee, *How Employers Tracking Your Health Can Cross the Line and Become Big Brother*, GUARDIAN (May 1, 2015, 8:30 AM), http://www.theguardian.com/lifeandstyle/us-money-blog/2015/may/01/employers-tracking-health-fitbit-apple-watch-big-brother.

B. The Evolving Nature of Privacy Harms Experienced by the Poor

While many Americans express unease over a perceived loss of privacy,[28] the harms to the poor from surveillance regimes reach far beyond generalized anxiety.[29] This is because many surveillance systems that surround the poor are purposefully designed to deliver a message of stigma to the subject while reinforcing societal stereotypes about dependency.[30] In turn, these stereotypes drive punitive laws directed at the poor.[31] Even if they are not always visible, privacy harms to the poor are real and can have physical and psychological impacts. For instance, surveillance can discourage the poor from accessing needed help or engaging with social and financial institutions due to fears associated with monitoring.[32] Moreover, welfare recipients suffer psychological injuries related to a loss of self-agency and reproach that can further trap them below the poverty line.[33] Similarly, in the low-wage workplace, invasive surveillance can result in disproportionate levels of psychological problems, including depression, which in turn may lower employee productivity and employer profits.[34]

---

28. On public opinion regarding privacy and its trade-offs, *see* Lee Rainie & Shiva Maniam, *Americans Feel the Tensions Between Privacy and Security Concerns*, PEW RES. CTR. (Feb. 19, 2016), http://www.pewresearch.org/fact-tank/2016/02/19/americans-feel-the-tensions-between-privacy-and-security-concerns/.

29. *See* Gilman, *supra* note 1, at 1394.

30. *See* Gustafson, *Degradation Ceremonies, supra* note 19 at 343–44, 354; Harry Murray, *Deniable Degradation: The Finger-Imaging of Welfare Recipients*, 15 SOC. F. 39, 40–42 (2000).

31. *See* Gustafson, *Degradation Ceremonies*, *supra* note 19, at 348; LOÏC WACQUANT, PUNISHING THE POOR: THE NEOLIBERAL GOVERNMENT OF SOCIAL INSECURITY 1–3 (2009).

32. *See* Sarah Brayne, *Surveillance and System Avoidance: Criminal Justice Contact and Institutional Attachment*, 79 AM. SOC. REV. 367 (2014); ROBERT MOFFITT ET AL., DEP'T OF HEALTH & HUMAN SERVS., A STUDY OF TANF NON-ENTRANTS: FINAL REPORT TO THE OFFICE OF THE ASSISTANT SECRETARY FOR PLANNING AND EVALUATION 2, 14 (2003) (reporting that new welfare report policies discourage participation).

33. *See* Nancy Goodban, *The Psychological Impact of Being on Welfare*, 59 SOC. SERV. REV. 403 (1985); GILLIOM, *supra* note 10, at 66–67, 78 (summarizing interviews with welfare recipients in Appalachia in the early 1990s); Nora Jacobson, *A Taxonomy of Dignity:A Grounded Theory Study*, 9 BMC INT'L HEALTH & HUM. RTS. 3, 7 (2009) (when the state treats marginalized people with a lack of dignity, the results can include "loss of respect, loss of self-worth, ego, sense of self, and soul, loss of status, social standing, and moral standing, loss of confidence and determination").

34. *See* David Holman, Claire Chissick & Peter Totterdell, *The Effects of Performance Monitoring on Emotional Labor and Well-Being in Call Centers*, 26 MOTIVATION & EMOTION 57, 74–79 (2002); Debora Jeske & Alecia M. Santuzzi, *Monitoring What and How: Psychological Implications of Electronic Performance Monitoring*, 30 NEW TECH., WORK & EMP. 62 (2015); M.J. Smith et al., *Employee Stress and Health Complaints in Jobs with and Without Electronic*

The class differential in privacy harms also extends to life online. In the early days of the Internet, the poor faced a stark digital divide—they were excluded from online life due to an inability to afford computers and broadband access.[35] These days, low-income Americans are increasingly online, often through the use of smartphones.[36] Mobile access has helped to narrow the digital divide, but has left low-income Americans vulnerable to new forms of tracking. Poor Americans are considerably less likely to use Apple phones, which provide more robust encryption and are generally less susceptible to being hacked compared to their less expensive Android counterparts.[37] Wealthy and higher-educated Americans are more likely to use iPhones, on which data is encrypted by default and is more difficult for police, government, or phone companies to intercept.[38]

Beyond access, researchers have brought attention to the digital literacy skills divide, which can also impact low-income Internet users' exposure to privacy- and security-related harms.[39] In addition, certain inequalities relate not to low-income Americans being more likely to experience a given harm, but to their propensity to face harsher consequences as a result of those harms, in part due to a lack of resources to seek redress. Consider identity theft, a growing concern shared across social classes. This crime is particularly devastating for low-income individuals, who face not only financial losses that impact their ability to meet basic needs such as housing and utility services, but are also left coping with more severe consequences of someone else using their identity, such as wrongful arrests, improper child support garnishments, and harassment by collection

---

Performance Monitoring, 23 APPLIED ERGONOMICS 17, 23–27 (1992); Scott C. D'Urso, *Who's Watching Us at Work? Toward a Structural-Perceptual Model of Electronic Monitoring and Surveillance in Organizations*, 16 COMM. THEORY 281, 287 (2006).

35. *See generally* Janet Thompson Jackson, *Capitalizing on Digital Entrepreneurship for Low-Income Residents and Communities*, 112 W. VA. L. REV. 187 (2009).

36. *See* Aaron Smith, *U.S. Smartphone Use in 2015*, PEW RES. CTR. (Apr. 1, 2015), http://www.pewInternet.org/2015/04/01/us-smartphone-use-in-2015/.

37. *See* Aaron Smith, *Smartphone Ownership–2013 Update*, PEW RES. CTR. (June 5, 2013), https://perma.cc/KE53-9R5X; ACLU Complaint, *Request for Investigation and Complaint for Injunctive Relief*, Apr. 16, 2013, *available at* https://perma.cc/W9AT-YUFD.

38. *See* Christopher Soghoian, *Your Smartphone is a Civil Rights Issue,* TED TALK (Oct. 2016), https://perma.cc/29ZQ-MFFB; Kaveh Waddell, *Encryption is a Luxury*, ATLANTIC (Mar. 28, 2016), https://perma.cc/DY7K-PTWY.

39. *See* Eszter Hargittai, *Second-Level Digital Divide: Differences in People's Online Skills*, 7 FIRST MONDAY (2002), https://perma.cc/A3FA-W9D2; Paul DiMaggio et al., *From Unequal Access to Differentiated Use: A Literature Review and Agenda for Research on Digital Inequality*, *in* SOCIAL INEQUALITY 355 (Kathryn Neckerman, ed., 2004).

agencies.[40]

It is also important to recognize that for the poor, overt and covert surveillance systems interact with one another. For instance, after the welfare system collects an applicant's data, this data is electronically shared and compared across multiple government and commercial databases in order to determine eligibility and ferret out fraud. These systems have the potential to make the application process easier and more streamlined for both social service offices and applicants.[41] At the same time, these databases are plagued with outdated, inaccurate, and incomplete data.[42] As a result, thousands of people have been denied benefits to which they would otherwise be entitled. [43]

On a day-to-day basis, welfare benefits and food stamps are distributed electronically and monitored to see how recipients are spending their money, thereby limiting "clients' autonomy, opportunity, and mobility: their ability to meet their needs in their own way."[44] Moreover, public benefits data is fed to law enforcement systems and vice-versa, in an ongoing loop of digital records sharing.[45] In addition, the personal data held in public benefits systems is at risk of security breaches. For instance, the Lifeline program, which provides wireless phones to low-income people, requires applicants to share personally identifiable information including income, social security numbers, and drivers' license numbers.[46] The communications industry has fought proposed government requirements that they keep this data secure.[47]

In sum, surveillance of the poor is broader, more invasive, and more difficult to redress than surveillance of other groups, and the overlap among government, commercial, and institutional data flows creates unique challenges for maintaining the accuracy and security of records.

---

40. *See* Sarah Dranoff, *Identity Theft: A Low-Income Issue*, 17 AM. BAR ASSOC. DIALOGUE MAG., Winter 2014, https://perma.cc/ES6W-QFZ8.

41. *See* STAN DORN & ELIZABETH LOWER-BASCH, CTR. FOR LAW & SOC. POLICY, MOVING TO 21ST-CENTURY PUBLIC BENEFITS: EMERGING OPTIONS, GREAT PROMISE, AND KEY CHALLENGES 4–6 (2012), https://perma.cc/SHM2-KX8L.

42. *See id*. at 15–19; Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1256–58 (2008).

43. *Id*. at 1256–57, 1268–73.

44. *See* Eubanks, *Technologies of Citizenship*, *supra* note 16, at 90–91.

45. *See* Kaaryn Gustafson, *The Criminalization of Poverty*, 99 J. CRIM L. & CRIMINOLOGY 643, 647, 667–71 (2009).

46. *See* Margaret Harding McGill, *CTIA Defends Its Challenge of FCC Lifeline Privacy Changes*, LAW 360 (Oct. 20, 2015, 6:02 PM), https://perma.cc/7KJ4-K28Y.

47. *See id*.

Low-income Americans live in communities with overt and omnipresent surveillance, and this oppression extends into the more covert surveillance that happens online. In both systems, the harms suffered by the poor can be concrete and stigmatizing.

C. Big Data Analytics, Social Media, and the Potential for Negative Impacts Among Low-Income Communities

Poor Americans face heightened risks from big data (that is, the collection, aggregation, analysis, and use of mass amounts of digital information gathered and shared about individuals).[48] Big data "gathers its contents from a myriad of online user interactions and infrastructure sensors, ranging from online transactions, search queries, and health records to communication networks, electric grids, and mobile phones."[49] Big data systems scoop up personal information when people shop in stores or online, visit websites, pay bills, use social media and mobile applications, or use devices such as fitness trackers.[50] Certain systems then combine that information with more "traditional" metrics that are used to evaluate individuals—such as credit history, criminal background records, and educational testing scores.[51] Big data holds tremendous promise to improve problem solving through greater insight into complex issues. It also raises the peril of information mischaracterization, misinterpretation, and abuse—all without the knowledge of the subjects whose data is being manipulated.[52]

The obfuscation of big data methods that now occurs across many industries has been variably described by scholars as creating a "black box society,"[53] a "transparency paradox,"[54] and a lack of "algorithmic

48. *See* Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 96 (2014).

49. *Id*. at 96.

50. FED. TRADE COMM'N, BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION? UNDERSTANDING THE ISSUES 1 (2016), https://perma.cc/5LAG-83MZ.

51. *See generally infra* case studies discussed in Part III.

52. *See generally* danah boyd & Kate Crawford, *Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon*, 15 INFO., COMM. & SOC'Y 662, 663 (2012), https://perma.cc/5BLE-KFBC ("Like other socio-technical phenomena, Big Data triggers both utopian and dystopian rhetoric.").

53. FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 3 (2015).

54. Neil M. Richards & Jonathan H. King, *Three Paradoxes of Big Data*, 66 STAN. L. REV. 41, 42 (2013).

accountability."[55] Limited public awareness about these practices has contributed to a regulatory environment in which the aggregation and brokering of personal data has largely gone unchecked.[56] This has renewed concerns about the impact of these information asymmetries on low-income populations, and the extent to which these dynamics may be contributing to increases in economic inequality in the United States.[57]

On the surface, big data collection may appear to be less stigmatizing for the poor than many of their other interpersonal interactions with the government because it is invisible. Algorithmic assessments that passively take one's online activities into account do not make the same kind of dehumanizing requests of low-income people (such as asking for urine samples or sexual history) as is sometimes required during the process of seeking public benefits.[58] Yet, the use of big data can injure the economic stability and civil rights of the poor, such as when they are targeted for predatory financial products, charged more for goods and services online, or profiled in ways that limit their employment and educational opportunities.[59] Conversely, big data can also result in the exclusion of marginalized groups from desirable opportunities "because they are less involved in the formal economy and its data-generating activities [or because they] have unequal access to and relatively less fluency in the technology necessary to engage online, or are less profitable customers or important constituents and therefore less interesting as targets of observation."[60]

Many poor people, and particularly low-income people of color, live in crowded urban environments that are under constant surveillance by law

---

55. ALEX ROSENBLAT, TAMARA KNEESE & DANAH BOYD, ALGORITHMIC ACCOUNTABILITY, DATA & SOC'Y RES. INST. (2014), http://ssrn.com/abstract=2535540.

56. *See* Alice E. Marwick, *How Your Data Are Being Deeply Mined*, N.Y. REV. BOOKS, Jan. 9, 2014, http://www.nybooks.com/articles/2014/01/09/how-your-data-are-being-deeply-mined/.

57. *See* Newman, *supra* note 3, at 2–3; FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 46–48 (2014), https://perma.cc/CAZ5-54PC; EXEC. OFFICE OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES 6–7 (2014), https://perma.cc/C6SB-BLC6 [hereinafter SEIZING OPPORTUNITIES]; Joseph W. Jerome, *Buying and Selling Privacy: Big Data's Different Burdens and Benefits*, 66 STAN. L. REV. ONLINE 47, 51 (2013) ("Most of the biggest concerns we have about big data—discrimination, profiling, tracking, exclusion—threaten the self-determination and personal autonomy of the poor more any other class.").

58. *See supra* note 23 and accompanying text.

59. *See* FED. TRADE COMM'N, *supra* note 50, at 9–11 (summarizing concerns raised by stakeholders).

60. *See* Barocas & Selbst, *supra* note 4, at 685.

enforcement.[61] As a result, they are much more likely than people in other contexts to become entangled with the criminal justice and child welfare systems, both of which are highly stigmatizing and privacy-stripping.[62] These interactions then become embedded in big data inputs and assessments, which in turn limits housing, employment, and educational opportunities for those affected.[63] At the same time, big data analysis of poor peoples' information is aggregated and used to craft policies and rules that then shape poor peoples' lives. As one commentator noted with regard to the collection of data from homeless people as a condition of receiving services, "whether or not a specific individual can be related back to data generated out of that individual, the life of that data will absorb and transform the life of that individual."[64]

The practice of specifically incorporating *social media data* into big data systems is becoming increasingly common in a wide range of industries. While thousands of data points—including both structured and unstructured data—may feed into various assessment tools, unique features of social media data can make it especially problematic for ensuring fairness and preventing bias in various forms of decision-making.[65] As boyd et al. argue, the value in analyzing social media data stems, in part, from the ability to assess both to whom you are connected and "who[m] you are like" based on your behavior and preferences. This information becomes valuable to a range of entities, from marketers to employers to law enforcement.[66]

Poor Americans have long suffered from guilt by association, meaning they bear the stereotypes and stigma of their social class (and race and gender) in ways that impede their economic progress and well-being.[67] As

61. *See* Kathryne M. Young & Joan Petersilia, *Keeping Track: Surveillance, Control, and the Expansion of the Carceral State*, 129 HARV. L. REV. 1318, 1322 (2016); MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS 124–26 (2013); LORI BETH WAY & RYAN PATTEN, HUNTING FOR "DIRTBAGS": WHY COPS OVER-POLICE THE POOR AND RACIAL MINORITIES 3, 5, 103, 116, 136 (2013).

62. On criminal justice, *see id*. On child welfare, *see generally* DOROTHY ROBERTS, SHATTERED BONDS: THE COLOR OF CHILD WELFARE 8 (2002).

63. *See infra* Part III. On the loss of opportunities wrought by possessing a criminal record, *see generally* JAMES B. JACOBS, THE ETERNAL CRIMINAL RECORD, 225–314 (2015).

64. Craig Willse, "*Universal Data Elements," or the Biopolitical Life of Homeless Populations*, 5 SURVEILLANCE & SOC. 227, 245 (2008).

65. *See* boyd et al., *supra* note 5, at 43–57.

66. *Id*.

67. *See, e.g.*, JOEL F. HANDLER & YEHESKEL HASENFELD, BLAME WELFARE, IGNORE POVERTY AND INEQUALITY 70 (2007); Heather E. Bullock, *Justifying Inequality: A Social Psychological Analysis of Beliefs About Poverty and the Poor*, *in* THE COLORS OF POVERTY: WHY RACIAL AND

scholars who study the surveillance of marginalized groups have noted, "networks of association are not random, and who we know online is affected by offline forms of residential, educational, and occupational segregation."[68] The case studies in this Article highlight concrete examples of how these guilt-by-association effects, when paired with the networked nature of social media data, may exacerbate big data-related harms for the poor.

Simply opting out from using social media and other digital technologies to avoid these risks is not an option in today's digital world, and may be impossible for some forms of surveillance.[69] Increasingly, institutions—such as schools, workplaces, and social service agencies—require engagement on certain platforms to get access to information and resources. Internet access has become an essential conduit for commerce, educational information, job opportunities, government services, and to maintain social connections to friends and family. Low-income adults recognize these benefits, with more than eight in ten saying the Internet has improved their ability to "learn new things" and majorities reporting that the Internet has made them better informed about products and services.[70] If low-income users were to opt out of using certain websites or applications due to privacy concerns, they would also lose the ability to access the myriad opportunities associated with engagement in online life.

II. SURVEY OF PRIVACY AND SECURITY CONCERNS OF LOW-INCOME INDIVIDUALS

Low-income individuals, then, face pervasive and disproportionate scrutiny in connection with government services, institutional involvement, and low-wage work, and face new and evolving challenges due to the advent of big data analytics and "guilt by association." But in addition to these, poor Americans' patterns of *technology use* and *privacy-relevant behaviors* expose them to greater risk than their wealthier counterparts. In Part II, we demonstrate these disparities empirically.

ETHNIC DISPARITIES PERSIST 52, 57 (Ann Chih Lin & David R. Harris eds., 2008).

68. Eubanks, *supra* note 1.

69. Sarah Kessler, *Think You Can Live Offline Without Being Tracked? Here's What It Takes*, FAST COMPANY (Oct. 15, 2013, 6:00 AM), https://perma.cc/Q5UQ-42UY.

70. Kristen Purcell & Lee Rainie, *Americans Feel Better Informed Thanks to the Internet*, PEW RES. CTR. 2 (DEC. 8, 2014), https://perma.cc/R3DC-JGE4.

A. Challenges in Demonstrating Harm and Need for Empirical Research Highlighting Unique Vulnerabilities of Low-Income Groups

Prior empirical studies have found that low-income Internet users are significantly more likely than higher-income users to report negative experiences connected to their online activity.[71] For instance, poorer Internet users are more likely to say they had an email or social media account compromised, and are more likely to report having their reputation damaged by online activity.[72] However, the reporting of privacy-related harms in surveys relies on respondents being *aware* of the negative impacts in question. In cases of big-data-related decision-making and discrimination, it is nearly impossible for respondents to know what personal or behavioral information may have factored into an unfavorable outcome.

Recent qualitative studies have focused on understanding what behaviors might be associated with privacy-related vulnerabilities among low-status users. For instance, researchers have suggested that "marginal Internet users" who rely on digital literacy organizations for training and access, may be more likely to engage in online behaviors that make them susceptible to potential privacy problems, such as being tracked with third-party cookies or unwittingly disclosing their information to fraudulent or predatory websites.[73] In addition, legal scholars have noted the need for analysis that examines whether or not low-status users face magnified privacy vulnerabilities due to knowledge gaps about privacy and security-related tools.[74]

We build upon this framework of understanding privacy-related vulnerabilities and provide new insights into the behaviors and attitudes of low-income Internet and social media users, which are of particular relevance to discussions of big-data-driven analysis. In the section that follows, we address the intersection of privacy-related vulnerabilities and socioeconomic status through an empirical examination of tech-related behaviors among low-income groups, using data from a new, nationally representative survey. And while income is the primary focus of our analysis, it is not the only indicator of a person's socioeconomic status (SES). For instance, the American Psychological Association broadly

71. Lee Rainie et al., *Anonymity, Privacy, and Security Online*, PEW RES. CTR. 24 (Sept. 15, 2013), https://perma.cc/JT98-FCZ5.

72. *Id*.

73. Gangadharan, *supra* at 6.

74. Urban & Hoofnagle, *supra* note 10, at 4.

defines socioeconomic status as "the social standing or class of an individual or group," which is "often measured as a combination of education, income and occupation."[75]

B. Survey Methods and Goals

In order to help address the question of how privacy and security experiences vary across different socioeconomic groups, the Digital Trust Foundation supported the fielding of a robust, nationally representative twenty-minute random-digit-dial telephone survey of 3000 American adults eighteen and older. The sample included an oversample of adults with annual household incomes of less than $40,000.[76] This survey, fielded in November and December of 2015, explores how low-income adults' concerns about privacy fit into the larger scope of concerns they have in everyday life, and contributes to a deeper understanding of their technology-related behaviors and beliefs.[77]

Among the key findings in this new survey are several that suggest greater privacy-related vulnerabilities among low-income Internet users as compared to higher income groups. While some of these differences may be influenced by multiple contributing factors (such as one's age and education level), examining variations through the lens of income helps to inform policy and technology design decisions that are specifically tailored for low-income groups. In particular, the discussion that follows highlights certain sociotechnical behaviors that, when combined with a lack of sufficient legal protections and rapidly evolving industry practices, may result in increased exposure to surveillance and big-data related

75. *See* AM. PSYCHOLOGICAL ASS'N REPORT OF THE APA TASK FORCE ON SOCIOECONOMIC STATUS (2007), http://www.apa.org/pi/ses/resources/publications/task-force-2006.pdf (also noting that various indicators are often overlapping; those who live with lower household incomes also tend to have lower education levels and work in low-wage jobs).

76. This level of annual income represents roughly 200% of the current federal poverty level for a household of three. *See 2015 Poverty Guideline*, ASPE (2015), https://perma.cc/7VBF-UGKR. Current estimates from the Census Bureau suggest the average household size in the U.S. is 2.54 people. *See Families and Living Arrangements*, U.S. CENSUS BUREAU, https://perma.cc/53JT-3XYC. The survey included a standard measure of an individual's household income that asks each participant to report their "total family income from all sources, before taxes" for the previous year. *Id*.

77. Co-author Mary Madden is also the Principal Investigator for this project. The main survey report is scheduled for publication in 2017 and will include a detailed discussion of methods and sample design. Mary Madden, *Privacy, Security and Digital Inequality: How Technology Experiences and Resources Vary Based on Socioeconomic Status and Race*, DATA & SOC'Y RESEARCH INST. (forthcoming 2017). See *infra* Appendix for a summary of methods used for the survey.

harms among low-status groups.[78]

C. Patterns of Mobile Internet Use Unique to Low-Income Populations

Echoing previous findings from a wide range of empirical studies, the survey results indicate that low-income Internet users who own smartphones are significantly more likely than higher income groups to say they "mostly go online" using their cell phone.[79] Even when considering the fact that low-income adults are less likely to be Internet users and less likely to own smartphones overall, the share of low-income adults who rely on their mobile devices as their primary source of Internet connectivity still exceeds that of higher income groups.

Overall, 39% of all Internet users who own a smartphone say that their cell phone is the primary way they go online. Another 41% say they mostly use some other device, and 20% report that they use their cell phone and other devices equally. However, the differences at either end of the income spectrum are stark; 63% of smartphone Internet users who live in households earning less than $20,000 per year say they mostly go online using their cell phone, compared with just 21% of those in households earning $100,000 or more per year.

The survey findings also suggest that age is an important indicator; looking at broader income groups (less than $40,000 per year vs. $40,000 per year or more), Internet users ages eighteen to twenty-nine who have a smartphone and are in the lower income bracket are more likely to report a reliance on cell phones for Internet access when compared with young adults living in higher income households (62% vs. 46%).[80] However, there is an even larger gap among lower income smartphone owners ages thirty to forty-nine, who are more than twice as likely as higher income adults of the same age to say they mostly go online using their phone (71% vs. 29%).

---

78. All differences noted between various comparison groups discussed throughout this section are statistically significant. The differences were evaluated with an independent Z-test for significance at the 95% confidence level.

79. Maeve Duggan & Aaron Smith, *Cell Internet Use 2013*, PEW RES. CTR. (Sept. 16, 2013), https://perma.cc/49NH-37NN.

80. These broader income categories (<$40K/$40K+) are used when analyzing detailed age groups in order to allow for a large enough sample to make valid comparisons.

D. Privacy and Security Vulnerabilities Associated with Reliance on Mobile Devices

Recent media coverage about the strong encryption available on Apple iPhones has largely overshadowed the many well-documented privacy vulnerabilities associated with mobile devices and applications.[81] While current versions of the Apple iOS operating system may make it difficult for law enforcement to access the contents of a locked phone, cell phone users are still subject to a wide range of mobile surveillance possibilities that they may be unaware of, including advertisers' cross-device tracking,[82] cell site simulators[83] and in-store tracking by retailers.[84] In particular, location-related data, when gathered from mobile devices over a period of time and tied to both online and offline behaviors, can reveal an incredibly intimate portrait of users' daily lives.

---

81. *See* Paarijaat Aditya, et al., *Brave New World: Privacy Risks for Mobile Users*, *in* PROCEEDINGS OF THE ACM MOBICOM WORKSHOP ON SECURITY AND PRIVACY IN MOBILE ENVIRONMENTS 7 (2014), https://perma.cc/3FQL-TXMP.

82. Michael Whitener, *Cookies Are So Yesterday; Cross-Device Tracking Is In—Some Tips*, PRIVACY ADVISOR (Jan. 27, 2015), https://perma.cc/5WAZ-824G.

83. *See* Kim Zetter, *Turns Out Police Stingray Spy Tools Can Indeed Record Calls*, WIRED MAG. (Oct. 28, 2015, 3:00 PM), https://perma.cc/7GJ8-UEH4.

84. *See* Press Release, Fed. Trade Comm'n, Retail Tracking Firm Settles FTC Charges It Misled Consumers About Opt Out Choices (Apr. 23, 2015), *available at* https://perma.cc/7H2P-9GSB.

Table 1. Mostly Mobile Internet Use by Income

Responses to the question: "Overall, when you use the Internet, do you do that mostly using your cell phone or mostly using some other device like a desktop, laptop or tablet computer?"

	Internet users who own a smartphone	Less than $20K	$20K-under $40K	$40K-under $75K	$75K-under $100K	$100K or more
	(a)	(b)	(c)	(d)	(e)	(f)
Mostly on cell phone	39%	63%[def]	58%[def]	34%[f]	24%	21%
Mostly on something else	41%	22%	28%	42%[bc]	61%[bcd]	52%[bc]
Both equally (Volunteered)	20%	14%	13%	24%[bc]	15%	26%[bce]

Note: Significant differences within rows are noted with superscript letters indicating the column to which the item should be compared. For instance, the 63% of those in households earning less than $20,000 per year who say they use the Internet mostly from their cell phone is significantly higher than those in households earning $40,000 or more. The differences noted here were evaluated with an independent Z-test for significance at the 95% confidence level. Column (a), which displays responses for all Internet users who own a smartphone, is not included in the significance testing. A small number of users (1% or less) offered a volunteered response of "It depends" or "Don't know" in response to this question; those responses are omitted.

Source: Privacy and Security Experience of Low-Socioeconomic Status Populations Survey, November 18–December 23, 2015, including an oversample of adults living in households earning less than $40,000 per year. Interviews were conducted in English and Spanish (Total n=3,000 US adults age 18 and older, n=1,724 for Internet users who own a smartphone).

A recent study published in *Nature* magazine illustrated that human mobility traces are a highly unique and sensitive form of data that pose considerable re-identification risks with only a handful of spatio-temporal data points. As the authors note:

> Mobility data is among the most sensitive data currently being collected. Mobility data contains the approximate whereabouts of individuals and can be used to reconstruct individuals' movements across space and time . . . . While in the past, mobility traces were only available to mobile phone carriers, the advent of smartphones and other means of data collection has made these broadly available.[85]

Beyond the broadcasting of location-related data, mobile applications create various vulnerabilities for smartphone users who rely on their phones as their primary mode of Internet access. Of particular relevance to debates about big data-related harms is the fact that mobile applications have not always offered consistent access to privacy policies or privacy controls for information sharing.[86] While certain mobile operating systems, such as iOS, offer simplified ways to manage location-sharing preferences within certain apps and tools to limit ad tracking, granular application settings to control other forms of in-app content sharing vary widely. In some cases—particularly with older versions of social media applications—a user must navigate to the website associated with a given app in order to change default settings.[87] Some applications automatically opt-in users to higher levels of sharing than they may be aware of or change the terms of service to retroactively apply to content that users had previously posted to the platform.[88]

---

85. Yves-Alexandre de Montjoye et al., *Unique in the Crowd: The Privacy Bounds of Human Mobility*, 3 SCI. REPS. 1 (2013) (noting that Apple's privacy policy allows the sharing of users' location data with "partners and licensees" and estimating that geo-location data for roughly half of all iOS and Android traffic is made available to advertising networks).

86. *See*, *e.g.*, FUTURE OF PRIVACY FORUM, MOBILE APPS STUDY (2012) (finding that just over half (53%) of the seventy-five paid apps reviewed for the study provided users access to a privacy policy).

87. This was the case with older versions of the Facebook mobile app, which did not have the inline audience selector. Changing the audience for the content shared on these older versions of the app requires users to navigate on the website to specific settings for "old versions of Facebook for mobile." *See How Do I Set the Audience When I'm Using an Older Version of Facebook for Mobile That Doesn't Have an Audience Selector?*, FACEBOOK, https://www.facebook.com/help/260276693997558/?ref=u2u.

88. One recent example that surprised social media users was Snapchat's retroactive changes

Many third-party mobile applications have been shown to access more data than is necessary for the application to function.[89] And some mobile applications have been shown to embed software that can surreptitiously perform other functions, such as monitoring a device's microphone without a user's permission.[90]

There have also been notable security-related vulnerabilities associated with the operating systems on mobile devices. In 2013, the ACLU filed a complaint with the FTC noting that many smartphone owners were using a version of the Android operating system that had "known, exploitable security vulnerabilities for which fixes have been published by Google, but have not been distributed to consumers' smartphones by the wireless carriers and their handset manufacturer partners."[91] These kinds of security vulnerabilities are likely to be especially acute for low-income groups, who are more likely to be "smartphone dependent" for all or most of their internet connectivity.[92]

E. Social Media Use, Privacy-Protective Behaviors, and Confidence in Skills

Lower income Internet users are modestly more likely than Internet users in higher income households to say they use social media such as Facebook, Twitter or Instagram: 81% of online adults in households earning less than $20,000 per year say they use social media, compared with 73% of online adults in households earning $20,000 or more.[93] Most of this difference is attributable to the relative youthfulness of lower income Internet users, as online adults under the age of fifty are equally likely to use social media, regardless of income.

However, focusing more closely on privacy-related behaviors *within* the population of social media users reveals several notable variations by

---

regarding their use of user images. *See* Sally French, *Snapchat's New 'Scary' Privacy Policy Has Left Users Outraged*, MARKETWATCH (Nov. 2, 2015, 4:13 PM). http://www.marketwatch.com/story/snapchats-new-scary-privacy-policy-has-left-users-outraged-2015-10–29.

89. *See, e.g.*, *What They Know–Mobile*, WALL ST. J. (APR. 3, 2017), https://perma.cc/LWK3-P7DC; *Apps Permissions in the Google Play Store*, PEW RES. CTR. (Nov. 10, 2015), https://perma.cc/TEU8-E28T.

90. *See* Press Release, Fed. Trade Comm'n, FTC Issues Warning Letters to App Developers Using 'Silverpush' Code (Mar. 17, 2016), *available at* https://perma.cc/N4YK-DH2Y.

91. *See* ACLU, *supra* note 37, at 1.

92. Smith, *supra* note 36.

93. This difference is modest, but statistically significant when comparing these broad groups.

income.[94] For instance, social media users in the lowest income bracket are significantly *less* likely than higher earning groups to say they have used privacy settings to restrict access to the content they post online—whether on social media platforms or other websites. Among social media users living in households earning less than $20,000 per year, 65% say they have used privacy settings to limit who can see what they post online, while 79% of those in wealthier households say they have done this.[95]

Some of these behaviors are also associated with lower levels of confidence in certain privacy-related skills and knowledge. Low-income social media users are less likely to feel as though they "know enough" about managing the privacy settings for the information they share online (65% vs. 77%) and are less likely to feel they have a good understanding of the privacy policies for the applications and websites they use (64% vs. 74%). At the same time, low-income social media users are more likely than higher earning groups to feel as though it would be "somewhat" or "very" difficult to find tools and strategies that would help them protect their personal information online (25% vs. 15%).

Low-income social media users are also less likely to engage in other privacy-protective strategies that may impact the way they are tracked online. For instance, they are less likely to say that they have avoided communicating online when they had sensitive information to share. About half (52%) report this, compared with 63% of social media users in wealthier households. Similarly, a smaller share of low-income social media users say they have set their browsers to turn off cookies or notify them before receiving a cookie (47% vs. 58%).

While using privacy settings, self-censoring communications, and restricting the use of cookies may limit some forms of tracking and profiling, the kinds of social media data input that are assessed by algorithmic systems can also include profile information that many users may not realize has retroactively become publicly available, or is made accessible to advertisers through third-party apps. For instance, over time, the information that has been made public by default on networks such as

---

94. Throughout this section, "low-income social media users" are defined as those living in households earning less than $20,000 per year. Comparisons are made with social media users living in households above that income threshold, and any differences included in the discussion are statistically significant.

95. The question about privacy settings is not limited to social media and could include the use of settings for other kinds of applications, platforms, and profiles.

Facebook has changed considerably.[96]

Table 2. Privacy Strategies Among Social Media Users by Income

The percentage who responded "yes" to the question: "While using the Internet, have you ever done any of the following things?"

	All social media users	Less than $20K	$20K or more
	(a)	(b)	(c)
Used privacy settings to limit who can see what you post online	76%	65%	79%[b]
Avoided communicating online when you had sensitive information to share	60%	52%	63%[b]
Set your browser to turn off cookies or notify you before you receive a cookie	56%	47%	58%[b]

Note: Significant differences within rows are noted with superscript letters indicating the column to which the item should be compared. The differences noted here were evaluated with an independent Z-test for significance at the 95% confidence level. Column (a), which displays responses for all social media users, is not included in the significance testing.

Source: Privacy and Security Experience of Low-Socioeconomic Status Populations Survey, November 18–December 23, 2015, including an oversample of adults living in households earning less than $40,000 per year. Interviews were conducted in English and Spanish. (Total n=3,000 US adults age 18 and older, n=1,613 for social media users).

As Hartzog et al. notes, the current big data landscape has made it increasingly difficult for users to effectively restrict access to their personal disclosures—even when they make well-intended efforts to do so.

96. Alessandro Acquisti, Laura Brandimarte & George Loewenstein, *Privacy and Human Behavior in the Age of Information*, 347 SCI. 513 (2015), https://perma.cc/THN5-BXSE.

> In the past, individuals have been able to roughly gauge whether aspects of their daily routines and personal disclosures of information would be safeguarded at any appropriate level of privacy protection by (sometimes implicitly) guessing the likelihood their information would be discovered or understood by third parties who have exploitative or undesirable interests. In the age of big data, however, the confidence level associated with privacy prognostication has decreased considerably, even when conscientious people exhibit due diligence.[97]

The resulting environment is one in which data brokers are able to glean a wide array of insights—such as usernames and friend connections—from social media activity. The FTC's *Data Brokers* report documents the various ways that social media data are being scraped from publicly available websites and combined with a wide range of other behavioral data to create and consumer profiles.[98] These kinds of practices can affect consumers across the socioeconomic spectrum, but low-income populations have been specifically targeted for their vulnerability. As a Senate Commerce Committee report on data broker practices identified, the poor have been profiled into various "financially vulnerable" market segments such as "Rural and Barely Making It," and "Fragile Families."[99] Such lists make it possible for marketers to easily target vulnerable consumers for dubious financial products such as payday loans, online classes, or debt relief services.[100]

The survey results indicate that low-income social media users also have a range of privacy-related concerns that are more pronounced when compared with higher income groups.[101] Most directly related to this

97. Woodrow Hartzog & Evan Selinger, *Big Data in Small Hands*, 66 STAN. L. REV. ONLINE 81, 84 (2013), https://perma.cc/S2G5-5NSH.

98. FED. TRADE COMM'N, DATA BROKERS, *supra* note 57, at 13–14.

99. MAJORITY STAFF OF OFFICE OF OVERSIGHT & INVESTIGATIONS, S. COMM. ON COMMERCE, SCI. & TRANSP., 113TH CONG., A REVIEW OF THE DATA BROKER INDUSTRY: COLLECTION, USE, AND SALE OF CONSUMER DATA FOR MARKETING PURPOSES 24–26 (2013) (staff report for Sen. Rockefeller IV, Chairman, S. Comm. of Commerce, Sci. & Transp.).

100. *See* AARON RIEKE ET AL., UPTURN, CIVIL RIGHTS, BIG DATA, AND OUR ALGORITHMIC FUTURE: A SEPTEMBER 2014 REPORT ON SOCIAL JUSTICE AND TECHNOLOGY 8 (2014), https://bigdata.fairness.io/wp-content/uploads/2015/04/2015-04-20-Civil-Rights-Big-Data-and-Our-Algorithmic-Future-v1.2.pdf; UPTURN, LED ASTRAY: ONLINE LEAD GENERATION AND PAYDAY LOANS (2015), https://perma.cc/UE5J-EGMV.

101. Throughout this section, "low-income social media users" will continue to be defined as those living in households earning less than $20,000 per year. Comparisons are made with social

discussion, low-income users express deeper worries about commercial data collection. About half (52%) say they are "very concerned" about not knowing what personal information is being collected about them by companies or how it is being used, compared with just over a third (37%) of more affluent social media users.

If the poor are subject to more data collection and surveillance across an array of institutional interactions in their daily lives, they may accordingly face greater burdens and vulnerabilities associated with the inaccuracy of records being fed into data broker profiles. In addition, if they are less likely to engage in certain privacy-protective behaviors online, this may further link their social media activity and network connections to data broker profiles in ways that result in negative outcomes associated with increasing reliance on algorithmic decision-making systems.

The section that follows illustrates three case studies of emerging practices in using publicly available social media data to inform critical decisions that can affect low-status individuals' economic mobility. Social media sites like Twitter, Reddit, Facebook, and Instagram are a rich trove of data, much of which is public. One marketer explained that "[i]f Big Data is the water pouring out of your faucet, then social media is the reservoir that stream comes from."[102] Facebook, for instance, partners with a variety of data-brokers, including Acxiom, DataLogix, BlueKai, Epsilon and Experian, to develop more detailed profiles of users combining on and offline information.[103] These data brokers are the largest in the world and already boast masses of data. (Acxiom claims to have data on 700 million people; Epsilon a file on every American household; Datalogix "more than $1 trillion" in offline purchase-based data.)[104] Combining these data sources with the information that Facebook has from online interactions has allowed the company to develop an ad-targeting system so sophisticated that it "could hypothetically serve soda ads to teenagers who recently purchased a soft drink at a convenience store, or diaper ads to parents who bought baby food at a department store."[105] In the next

media users living in households above that income threshold. Mary Madden, *supra* note 77.

102. Dennis Hung, *The Impact of Big Data on Social Media Marketing Strategies*, TECH. CO (Jan. 22, 2016, 11:03 AM), https://tech.co/impact-big-data-social-media-marketing-strategies-2016-01.

103. Marketing Partners Directory, FACEBOOK, https://perma.cc/G9EQ-845J (last visited Apr. 2017); Alex Senemar, *Facebook Partners with Shadowy 'Data Brokers' to Farm Your Information*, SHERBIT (Apr. 25, 2016), https://perma.cc/5MJ2-63FF.

104. Senemar, *supra* note 103. Also note that Datalogix is owned by Oracle Corporation. *See Oracle and Datalogix*, https://perma.cc/YR3R-82VK (last visited June 3, 2017).

105. Senemar, *supra* note 103.

section, we provide three examples of how social media data use by big data systems could contribute to inequality and differentially harm low-income individuals.

III. CASE STUDIES AND LEGAL ANALYSIS

A. Employment

1. The Use of Social Media to Determine Employability

The use of automated assessment methods to determine "employability" among job candidates has become a desirable feature of current Applicant Tracking Systems (ATS). ATS software is designed to simplify the hiring process and automate the review of resumes and applications for employers, who sometimes face the daunting task of sifting through thousands (or even millions) of applicants. "Using highly granular data about workers' behavior both on and off the job, entrepreneurs are building models that they claim can predict future job performance."[106] As early as 2012, industry experts claimed that the vast majority of Fortune 500 companies were using some kind of ATS system to screen candidates.[107] And a recent article from *HR Today* suggests that integrating the screening of social media profiles in Applicant Tracking Systems is among the top trends in the field: "By integrating the recruiting platform with such sites as LinkedIn, Facebook, and Twitter, recruiting teams can post job openings to a worldwide audience, at the same time mining a potential applicant's personal profile on social networks for deeper insights."[108]

The insights gleaned from social media can serve to weed out candidates who, for a variety of reasons, may not be seen as the best fit for the job.[109] These assessments can be made from a range of social media

106. Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857, 860 (2017).

107. *See* Lauren Weber, *Your Résumé vs. Oblivion: Inundated Companies Resort to Software to Sift Job Applications for Right Skills,* WALL ST. J. (Jan. 4, 2012), https://perma.cc/U3LD-JVPB.

108. Russ Banham, *2016 Trends in Applicant Tracking Systems*, HR TODAY (Feb. 2, 2016), https://perma.cc/3C3T-D2AD.

109. These practices are not limited to the U.S. and have created a secondary market for tools targeted at job candidates, such as the UK-based "Social Score" that allows applicants who have been rejected for a job to "see what employers see" when conducting a social media check. MY SOCIAL

data, including content analysis of social media posts, assessments of personality type from likes on a social media profile, or analysis of a potential employee's network connections to measure their "social capital" within a certain field.[110] As one CEO of a recruitment solutions provider notes, "[t]he knowledge and data acquired from online social practices allows recruiters to analyze the successes and shortcomings of candidates for greater relationship building."[111] In some cases, companies that provide ATS software offer social media background checks as a separate service to highlight a candidate's shortcomings. For instance, InfoCheckUSA offers a "Not FCRA compliant" social media background check report for $24.95 that will help employers "see what kind of person you are dealing with" and will identify activities such as "[e]xcessive Twittering or social media activity while on the clock."[112]

Some companies that provide applicant tracking solutions and predictive analytics platforms have offered job seekers advice on navigating the new world of social media assessments as part of the hiring process. HireVue, a company that promises to help employers screen "200% more candidates, land 13% more top performers, reduce poor performer hiring by 17%, and drive turnover down by 28%,"[113] published this note of caution for job applicants:

> Social media can be tricky with their privacy settings. Make sure you read through every setting and what it can and cannot filter for you . . . . [P]rofile pictures are public regardless of your profile page being private or public, so be careful with the content you post and make sure you understand how your posts are viewed. Social media can be your greatest asset or your biggest failure.[114]

While the exact variables that factor into an assessment are difficult to uncover, some companies provide descriptions of their evaluation process on their websites. One such company, Social Intelligence, argues that employers should use their service to "[a]void legal restrictions by

---

SCORE, https://perma.cc/YAX6-349C.

110. *See* Cathy O'Neil, *How Algorithms Rule our Working Lives*, GUARDIAN (Sept. 1, 2016, 1:00 PM), https://www.theguardian.com/science/2016/sep/01/how-algorithms-rule-our-working-lives.

111. Banham, *supra* note 108 (internal citation omitted).

112. *See Social Media Background Check*, INFO CHECK USA (Apr. 2017), https://perma.cc/7NBX-QAE3 (last visited June 3, 2017).

113. *See Maximum Recruiter Revenue*, HIRE VUE, https://web.archive.org/web/20150911005104/http://www.hirevue.com/recruiting/.

114. Emily Hatch, *Applying for a Job? Don't Be Another Social Media Failure!*, HIRE VUE (Sept. 25, 2014), https://perma.cc/5BAK-7V5R.

entrusting your social media screening with trusted online private investigators."[115] The company claims to be the only social media screening company that has an endorsement from the Federal Trade Commission (FTC) and employs Fair Credit Reporting Act (FCRA)[116]-certified analysts to identify "negative behavior" through a "cyber investigation" that can help companies to avoid "accusations of discrimination."[117] FCRA is the federal statute governing credit reporting. However, the categories of negative content that are highlighted through the exemplary review on their website are hardly straightforward; for instance, it is unclear what kinds of statements constitute "potentially unlawful" behavior, "potentially violent" posts, or what qualifies as "racism and/or demonstrations of intolerance" or "sexually explicit material." It appears to be neither a fully objective nor subjective process, but is presented as resulting in "accurate and dependable insight."[118]

Leaders in the HR profession have been engaged in an ongoing debate about the ethics of social media monitoring in various forms, but many companies are forging ahead with various workarounds.[119] In some cases, recruiters are hiring self-described FCRA-compliant companies to perform social media background checks on their behalf.[120] Researchers seeking to understand what filters are used to raise red flags about job candidates found that these indicators can include broad categories such as "At Risk Populations," "Potentially Unlawful Activity," and "Potentially Violent Behavior" (which includes the sub-filter of "Potentially aggressive verbiage").[121]

While these filters may unfairly exclude many applicants due to the misinterpretation or miscategorization of the content of their social media posts or photos, experimental research suggests that some employers may exclude applicants based solely on descriptive profile information such as

115. *Products*, SOCIAL INTELLIGENCE (Apr. 2017), https://perma.cc/GK7Y-4JEN.

116. For a detailed discussion of FCRA and its application to hiring, see *infra* Part III(A)(2)(a).

117. *Id*.

118. *Products*, SOCIAL INTELLIGENCE, https://perma.cc/GK7Y-4JEN (last visited June 22, 2017).

119. Johnathan A. Segal & Joyce LeMay, *Point/Counterpoint: Should Employers Use Social Media to Screen Job Applicants?*, SOC'Y FOR HUM. RESOURCE MGMT. (Nov. 1, 2014), https://perma.cc/DL7Q-4LMZ ("After an applicant has been interviewed, his or her membership in many protected groups is already known. So, checking his or her LinkedIn profile or Twitter handle is not likely to reveal much more than HR already knows.").

120. *See generally* ALEX ROSENBLAT, TAMARA KNEESE & DANAH BOYD, DATA & SOC'Y RESEARCH INST., NETWORKED EMPLOYMENT DISCRIMINATION (2014), https://perma.cc/279L-68TC.

121. *Id*.

religion.[122] Network information—which is made public by default across many social media sites—can also be used in problematic ways. While knowing who someone is connected to may provide valuable sourcing information to recruiters, it has the potential to create new forms of "networked discrimination" that may fall outside of current legal regulations, discussed below.[123] Overall, algorithms used in ATS systems and related social media screening services can harm job applicants when they contain inaccurate data about individuals, when their underlying statistical models are inaccurate, or when the data outcomes reflect pre-existing structural disadvantage.[124] And while antidiscrimination law does not currently restrict economic sorting based on personality, habits, and character traits, all of these indicators can be revealed through mobile devices and social media activity.[125] In addition, because low-income social media users are more likely than higher income users to post content publicly, less likely to feel they have a good understanding of privacy policies, and less likely to engage in certain protective strategies, they may inadvertently be subject to a greater range of harms when evaluated through the use of certain ATS tools.

2. Legal Analysis of Applicant Tracking Systems

There is little legal recourse from inaccurate or discriminatory employment screening reports due to gaps in existing laws. As boyd et al. have argued, there are currently no restrictions in place to protect against discrimination on the basis of one's personal network, despite the fact that our laws ban discrimination on the basis of race, color, sex, national origin, and other protected classifications.[126] Increasingly, algorithmic

122. ALESSANDRO ACQUISTI & CHRISTINA FONG, AN EXPERIMENT IN HIRING DISCRIMINATION VIA ONLINE SOCIAL NETWORKS (2015), http://ssrn.com/abstract=2031979 (reporting that when researchers created two fake social media profiles indicating religious affiliation and submitted job applications on their behalf to over 4000 employers, the Muslim candidate received a 13% lower callback rate compared to the Christian candidate.).

123. Rosenblat, et al., *supra* note 120.

124. Kim, *supra* note 106, at 874–84 (noting that data analytics can also be used for intentional discrimination, but that this danger is low given that employers do not need complex algorithms to discriminate on the basis of highly salient characteristics).

125. Peppet, *supra* note 25, at 125 (describing the ways in which the Internet of Things—physical devices embedded with digital data collection, such as fitness trackers and wireless-connected thermostat controls—may further magnify the proliferation of data streams being used to assess consumers' potential value and risk).

126. danah boyd, Karen Levy & Alice Marwick, *The Networked Nature of Algorithmic Discrimination*, in DATA & DISCRIMINATION:COLLECTED ESSAYS 54 (Seeta Pena Gangadharan &

means of decision-making provide new mechanisms through which discrimination may occur.[127]

The following section analyzes the major laws governing Applicant Tracking Systems: the Fair Credit Reporting Act (FCRA),[128] which aims to ensure accurate credit reports, and Title VII,[129] which prohibits employment discrimination. The bottom line is that there are scant legal limits on commercial data collection, and constraints on uses of data outputs are minimal. Our legal system largely relies on individuals to police their own privacy. Since low-income people are less likely to have confidence in and use privacy settings, they are especially vulnerable to discriminatory uses of big data by employers.

a. Fair Credit Reporting Act

Applicant Tracking Systems qualify as consumer reporting agencies (CRAs), which are regulated by FCRA. FCRA was enacted in 1970 to promote the accuracy, fairness, and privacy of personal information gathered by CRAs.[130] A CRA is an entity that assembles and generates consumer reports, which contain information "bearing on a consumer's . . . character, general reputation, personal characteristics, or mode of living" to determine the consumer's eligibility for employment, among other purposes.[131] CRAs must use "reasonable procedures to assure maximum possible accuracy of the information concerning the individual about whom the report relates."[132] The FCRA covers CRAs that use social media information in their reports,[133] which employment background reports

Virginia Eubanks, eds. 2014), https://www.newamerica.org/oti/policy-papers/data-and-discrimination/. If the social network were used as a proxy to discriminate against a protected class such as race, this would violate the law, but it would be very hard to prove. *See* Kim *supra* note 106, at 884.

127. *See* boyd, *supra* note 52, at 664; Kim *supra* note 106, at 9–10; Barocas & Selbst, *supra* note 4, at 674.

128. 15 U.S.C. §§ 1681 (2012).

129. *See generally* Civil Rights Act of 1964 § tit. VII, 42 U.S.C. § 2000e (1964).

130. *See generally* 15 U.S.C. § 1681.

131. 15 U.S.C. § 1681a(d)(1).

132. 15 U.S.C. § 1681e(b).

133. On May 9, 2011, the FTC issued a letter to a company called Social Intelligence Corporation, which conducts social media background screening for employers. In the letter, the FTC clarified that the company is a CRA because it "assembles or evaluates consumer report information that is furnished to third parties that use such information as a factor for establishing a consumer's eligibility for employment." Letter from Maneesha Mithal, Assoc. Dir., Fed. Trade Comm'n, to Renee Jackson, Counsel, Soc. Intelligence Corp. (May 9, 2011), https://www.ftc.gov/sites/default/files/documents/closing_letters/social-intelligence-corporation/110509socialintelligenceletter.pdf. In 2012,

increasingly contain.[134]

For their part, employers that use consumer reports must provide job applicants with notice and obtain applicants' written consent.[135] Of course, in today's job market, consent is mostly a formality, as few prospective employees are in a position to withhold it. If the employer then uses the report to take "adverse action" against the applicant (such as a failure to hire), the employer must notify the applicant and provide him or her with a copy of the credit report and a written summary of applicant's rights.[136] The applicant then has a short time period to identify and dispute any errors in the report, and upon expiration of that deadline, the employer can take the adverse action.[137] Notably, an employer faces no FCRA liability for failure to hire, whether based on an accurate or inaccurate report.[138]

In addition, FCRA leaves several other notable gaps. To begin with, it does not cover employers who gather their own information and conduct their own background checks, such as by checking prospective employees' social media accounts or public records accessible on the Internet.[139] Some estimates are that one-fifth to one-quarter of employers research job applicants themselves, using social networks and search engines.[140]

---

the FTC came to the same conclusion with regard to Spokeo, a data broker that used social media and other data to create detailed personal profiles of consumers, including information such as "hobbies, ethnicity, religion, participation on social networking sites, and photos." Press Release, Fed. Trade Comm'n, Spokeo to Pay $800,000 to Settle FTC Charges Company Allegedly Marketed Information to Employers and Recruiters in Violation of FCRA (June 12, 2012), https://perma.cc/WXG2-BBM9. Spokeo paid a $800,000 fine to resolve allegations that it violated FCRA "by failing to make sure that the information it sold would be used only for legally permissible purposes; failing to ensure the information was accurate; and failing to tell users of its consumer reports about their obligation under the FCRA." *Id*. Spokeo also posted fake endorsements of its products. *Id*.

134. Credit reports used in lending are also increasingly incorporating social network data, along with "exchanged messages, tagged photos, browsing habits, education, searches, and geo-spatial data from mobile phones." Nizan Geslevich Packin & Yafit Lev-Aretz, *On Social Credit and the Right to be Unnetworked*, 2016 COLUM. BUS. REV. 339, 344 (2016).

135. 15 U.S.C. § 1681b(b)(2) (2012) (this disclosure and authorization must be in a separate document from the employment application).

136. 15 U.S.C. § 1681b(b)(3).

137. 15 U.S.C. § 1681m(b)(2) (2012).

138. Kim, *supra* note 106, at 900 ("Thus, fair information practice principles are unlikely to significantly limit employer use of data models.").

139. *See* 15 U.S.C. § 1681a(d)(2)(A)(i) (2012); *see also* Peppet, *supra* note 25, at 128; Amy Schmitz, *Secret Consumer Scores and Segmentations: Separating Consumer "Haves" from "Have-Nots,"* 2014 MICH. ST. L. REV. 1411, 1426 (2014).

140. Alexander Reicher, *The Background of Our Being: Internet Background Checks in the Hiring Process*, 28 BERKLEY TECH. L.J. 115, 116 (2013). The FCRA similarly excludes information about "transactions or experiences between the consumer and the person making the report." 15 U.S.C. § 1681a(d)(2)(A)(i) (such as information provided by the applicant through a drug test or breathalyzer

Furthermore, many furnishers (entities that provide information to CRAs) and CRAs evade the FCRA by claiming not to be engaged in consumer reporting.[141]

The FTC, which is the primary enforcer of FCRA and privacy law in general, is sensitive to the big data risks faced by low-income consumers and job applicants, but does not have the staff or budget to investigate all these companies.[142] The Consumer Financial Protection Board (CFPB) was recently empowered by Congress to enforce FCRA and has flexed its muscles, particularly with regard to the legal responsibilities of furnishers.[143] Still, the CFPB is a controversial agency, and its enforcement agenda will likely depend on the prevailing political winds.[144] The agency is currently facing a constitutional challenge regarding its structure.[145]

The cases the FTC has successfully resolved demonstrate that even

---

test).

141. Kimani Paul-Emile, *Beyond Title VII: Rethinking Race, Ex-Offender Status, and Employment Discrimination in the Information Age*, 100 VA. L. REV. 893, 917 (2014).

142. Daniel J. Solove & Woodrow Harzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 600–02 (2014) [hereinafter *The New Common Law*]. Under Section 5 of the FTC Act, a one-hundred-year-old consumer protection law that far predates the rise of the Internet, the FTC can challenge businesses that engage in "unfair or deceptive acts or practices in or affecting commerce." Federal Trade Commission Act of 1914 § 5, 15 U.S.C. § 45(a)(1) (2012). For instance, under this authority, the FTC has gone after companies that violated their posted privacy policies and companies that altered privacy policies without consumer consent, as well as companies that failed to use reasonable and appropriate security practices to safeguard personal information. Solove & Hartzog, *The New Common Law*, *supra* 142, at 628–43 (cataloguing FTC actions finding unfair and deceptive practices). Several of these actions have resulted in consent orders, which have created a form of common law setting forth best practices regarding data use. *Id*. at 607. However, the FTC's resource limitations and its cautious approach to regulating consumer data privacy make alternate vehicles of enforcement essential. *See* Woodrow Hartzog & Daniel Solove, *The Scope and Potential of FTC Data Protection*, 83 GEO. WASH. L. REV. 2230 (2015). Overall, "the FTC's role is largely to discourage bad behavior, not to compensate affected parties." *Id*. at 2294.

143. Andrew M. Smith & Peter Gilbert, *Fair Credit Reporting Act and Financial Update–2015*, 71 BUS. LAW 661, 664–67 (2016). The CFPB was created by the Dodd-Frank Wall Street Reform and Consumer Protection Act, 12 U.S.C. § 5511(a) (2012), in response to the financial crisis of 2008 and officially became operational in 2011. *See generally* Leonard J. Kennedy, Patricia A. McCoy & Ethan Bernstein, *The Consumer Financial Protection Bureau: Financial Regulation for the Twenty-First Century*, 97 CORNELL L. REV. 1141 (2012) (giving account of CFB's history and scope of authority).

144. Adam J. Levitin, *The Consumer Financial Protection Bureau: An Introduction*, 32 REV. BANKING & FIN. L. 321, 341–42, 364–69 (20113).

145. PHH Corp. v. Consumer Fin. Prot. Bureau, 839 F.3d 1 (D.C. Cir. 2016) (holding the agency's structure unconstitutional and making the Director removable by the President at will). This judgment was later vacated in PHH Corp. v. Consumer Fin. Prot. Bureau, No. 15-1177, 2017 U.S. App. LEXIS 2733 (D.C. Cir. Feb. 16, 2017), and the case was heard en banc on May 24, 2017.

supposedly reputable CRAs often violate the law. For instance, the FTC sued HireRight, a company that provides criminal background checks to large companies like Monster and Oracle, for violations such as failing to provide adverse action notices, failing to conduct investigations of disputed information, and inaccurate reporting.[146] HireRight settled the case for $2.6 million.[147] Likewise, the CFPB brought enforcement actions against two of the largest employment background screening providers, which generate more than ten million reports a year, because the reports contained impermissible information and widespread inaccuracies.[148] For instance, the reports "included criminal records attached to the wrong consumers, dismissed and expunged records, and misdemeanors reported as felony convictions."[149] Under the consent order, the companies agreed to pay $10.5 million in damages to consumers and a penalty of $2.5 million.[150] Such blatant violations suggest that these practices may be widespread within the industry.

Another stumbling block to enforcement is that most people do not know what information CRAs are reporting about them.[151] Furthermore, it is nearly impossible to know how credit scoring algorithms work because data brokers consider this information a trade secret.[152] While data analytics is touted for its ability to reduce human biases, it often merely replicates them.[153] "Relying on data models instead of human decision-making is unlikely to counter structural forms of bias, because these models take existing workplace structures as givens."[154] Algorithms can unwittingly import biases encoded by software engineers without any

146. Press Release, Fed. Trade Comm'n, Employment Background Screening Company to Pay $2.6 Million Penalty for Multiple Violations of the Fair Credit Reporting Act: FTC Charges HireRight Solutions Incorrectly Listed Criminal Convictions on Reports of Some Consumers (Aug. 8, 2012), https://perma.cc/8MR4-AMMM.

147. *Id*.

148. Press Release, Consumer Fin. Prot. Bureau, CFPB Takes Action Against Two of the Largest Employment Background Screening Report Providers for Serious Inaccuracies (Oct. 29, 2015), https://perma.cc/3AQG-KZYE.

149. *Id*.

150. *Id*.; Consent Order, Consumer Protection Financial Bureau, In the Matter of General Information Services, Inc. and e-Background Checks.com, Inc., File No. 2015-CFPB-0028 (2015).

151. *See* Paul-Emile, *supra* note 141, at 926.

152. Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 5 (2014).

153. Kim, *supra* note 106, at 865.

154. *Id*. at 16. ("[But] data can be a useful tool for *diagnosing* both cognitive and structural forms of bias").

outside check on that process.[155] Perhaps most importantly for the case of big data analytics, while the FCRA is aimed at ensuring accurate information, it does not protect job applicants from inaccurate *inferences* that are drawn from that information.[156] In addition, the assessments covered by the FCRA are limited to determinations made about an individual; however, many data-driven scoring tools skirt these boundaries by making household-level assessments.[157]

Even if a consumer is aware that she has been denied employment due to inaccurate information, the process to correct that information is ineffective for most and favors those who are wealthy and well-connected.[158] Consumers must use an online system that typically results in a form response,[159] a system some have referred to as a "Kafkaesque no man's land,"[160] that more often than not fails to resolve the problem.[161]

This is troubling given the high error rates in credit reports.[162] The FTC reported that one in five credit reports contains errors, and overall, 5% of reports have errors that could result in a denial of credit.[163] Error rates are similarly high with regard to CRAs' criminal history reporting, which is plagued with false positive and false negative identifications, the reporting

155. *Id*. at 14; Citron & Pasquale, *supra* note 152, at 14 ("Credit bureaus may be laundering discrimination into black-boxed scores, which are immune from scrutiny.") (footnote omitted).

156. Peppet, *supra* note 25, at 128.

157. ROBINSON + YU, KNOWING THE SCORE: NEW DATA, UNDERWRITING, AND MARKETING IN THE CONSUMER CREDIT MARKETPLACE 2 (2014), https://perma.cc/8B4H-76TT ("To avoid regulatory limits, credit bureaus sell slightly aggregated information, such as the financial circumstances of a household, rather than an individual. This data can be used to target products to groups of consumers with great precision, based on the financial health of their household or neighborhood.").

158. Tara Siegel Bernard, *Credit Error? It Pays to Be on V.I.P. List*, N.Y. TIMES, May 14, 2011, *available at* http://www.nytimes.com/2011/05/15/your-money/credit-scores/15credit.html.

159. Chi Chi Wu, *Automated Injustice: How a Mechanized Dispute System Frustrates Consumers Seeking to Fix Errors in Their Credit Reports*, 14 N.C. BANKING INST. 139, 157-61 (2010).

160. Bernard, *supra* note 158.

161. FED. TRADE COMM'N, REPORT TO CONGRESS UNDER SECTION 319 OF THE FAIR AND ACCURATE CREDIT TRANSACTIONS ACT OF 2003 21 (2015), https://www.ftc.gov/system/files/documents/reports/section-319-fair-accurate-credit-transactions-act-2003-sixth-interim-final-report-federal-trade/150121factareport.pdf.

162. Maureen Mahoney, *Errors and Gotchas: How Credit Report Errors and Unreliable Credit Scores Hurt Consumers*, CONSUMERS UNION 7–12 (Apr. 9, 2014), https://perma.cc/6PBP-PPWY (describing errors due to mining together records for different people who share similar identification, stale information, inaccurate information, and identity theft).

163. FED. TRADE COMM'N, REPORT TO CONGRESS UNDER SECTION 319 OF THE FAIR AND ACCURATE CREDIT TRANSACTIONS ACT OF 2003 iv-vi (2012), https://www.ftc.gov/sites/default/files/documents/reports/section-319-fair-and-accurate-credit-transactions-act-2003-fifth-interim-federal-trade-commission/130211factareport.pdf.

of expunged and sealed information, and other similar errors.[164]

Furthermore, litigation brought against CRAs to challenge the accuracy of their reporting faces an uphill struggle, given that courts have interpreted the accuracy provision in FCRA as requiring "only that the consumer reporting agency must follow *reasonable procedures* to assure" accuracy.[165] In short, the burden on CRAs to be accurate is minimal.[166] Even the most accurate consumer report could contain information gathered about an individual from social media posts by themselves and their networks. This information may be "accurately" reported by a CRA, but subject to devastating and incorrect inferences by employers.

Private enforcement is barred altogether for certain FCRA protections. For instance, private litigants cannot sue furnishers for reporting inaccurate information about them.[167] Rather, consumers can only sue a furnisher for failing to conduct an investigation.[168] In addition, Congress barred private rights of action to enforce employer (and other end users of consumer reports) obligations to provide adverse action notices.[169]

---

164. Paul-Emile, *supra* note 141, at 908. On the range of data inaccuracies associated with information technology, *see* Elizabeth D. De Armond, *Frothy Chaos: Modern Data Warehousing and Old-Fashioned Defamation*, 41 VAL. U. L. REV. 1061, 1075–89 (2007).

165. Koropoulos v. Credit Bureau, Inc., 734 F.2d 37, 42 (D.C. Cir. 1984). For a summary of the range of court interpretations on the "maximum possible accuracy requirement," *see* Elizabeth Doyle O'Brien, Comment, *Minimizing the Risk of the Undeserved Scarlet Letter: An Urgent Call to Amend § 1681E(B) of the Fair Credit Reporting Act*, 57 CATH. U. L. REV. 1217, 1227–36 (2008).

166. Lawsuits challenging companies' technical violations, such as a failure to provide separate disclosure and authorization forms, have fared much better. *See* Roy Maurer, *Know Before You Hire: 2016 Employment Screening Trends*, SOC'Y FOR HUM. RESOURCE MGMT. (Jan. 20, 2016), https://www.shrm.org/resourcesandtools/hr-topics/talent-acquisition/pages/2016-employment-screening-trends.aspx. According to reports, class actions settled in 2015 included cases against BMW, Calvin Klein, Chuck E. Cheese, Food Lion, Home Depot, and Whole Foods, with settlements ranging from $716,000 to $3 million. *See* David N. Anthony & Julie D. Hoffmeister, *The Fair Credit Reporting Act: Not Just About Credit*, BUS. L. TODAY (June 2016), https://www.americanbar.org/publications/blt/2016/06/13_anthony.html; Thomas Ahearn, *Class Action Lawsuits Will Continue to Increase in Target Rich Background Screening Environment in 2016*, EMPLOYMENT SCREENING RESOURCES, Dec. 28, 2015, http://www.esrcheck.com/wordpress/2015/12/28/fcra-lawsuits-will-continue-to-increase-in-target-rich-background-screening-environment-in-2016/.

167. 15 U.S.C. §§ 1681s-2(c)(1), (d) (2012); Hopson v. Chase Home Fin. LLC, 14 F. Supp. 3d 774, 789–90 (S.D. Miss. 2014); Stafford v. Cross Country Bank, 262 F. Supp. 2d 776, 782–83 (W.D. Ky. 2003) (holding that furnishers can be sued by private consumers only for failure to properly investigate a consumer's dispute transmitted by a CRA).

168. 15 U.S.C. § 1681s-2(b).

169. 15 U.S.C. § 1681m(h)(8) (an employer is required to notify a rejected applicant after taking the adverse action and the notice should contain contact information of the CRA informing the applicant of their rights to dispute the accuracy of the report). On the barriers to private enforcement, *see* Alexandra P. Everhart Sickler, *The (Un)fair Credit Reporting Act*, 28 LOY. CONSUMER L. REV. 238, 256, 265–80 (2016).

Furthermore, state tort law does not fill the gap because the FCRA preempts most state law claims.[170] Thus, a consumer has limited recourse for a FCRA violation tied to use of an ATS report.[171]

b. Employment Discrimination Law

Under federal law, employers cannot discriminate on the basis of race, color, religion, sex, national origin, disability, age (over forty), genetic information, or military service.[172] If an employer uses an ATS report to draw inferences about an applicant's protected characteristics and then denies them employment, it might be charged with disparate treatment.[173] Or, if the algorithms underlying ATS generate hiring recommendations that disfavor protected groups, applicants might have a claim for disparate impact.[174] These are the two main forms of legally cognizable employment discrimination.[175] Yet as explained below, the algorithmic nature of the data mining that underlies ATS is unlikely to result in a successful lawsuit under either theory[176]—even though big data can "reproduce existing patterns of discrimination, inherit the prejudice of prior decision-makers, or simply reflect the widespread biases that persist in society."[177]

---

170. 15 U.S.C. 1681h(e). *See generally* Elizabeth D. De Armond, *A Dearth of Remedies*, 113 PENN. ST. L. REV. 1, 7–9 (2008).

171. Class action plaintiffs' lawyers have had success suing companies for technical violations of FCRA, unrelated to the accuracy of information, such as failure to use a standalone disclosure form. *See* Anthony & Hoffmeister, *supra* note 166, at 2. *But see* Just v. Target Corp., 187 F. Supp. 3d 1064 (D. Minn. 2016) (rejecting FCRA class action because employer conduct was not willful). However, these suits may falter in light of the Supreme Court's recent opinion in *Spokeo v. Robins,* holding that plaintiffs must establish an injury in fact that is both concrete and particularized. 136 S. Ct. 1540 (2016). A pure statutory violation with no other impacts may not meet this standard.

172. 42 U.S.C. §2000e et seq. For an overview of Title VII protections, *see* Robert Sprague, *Welcome to the Machine: Privacy and Workplace Implications of Predictive Analytics*, 21 RICH. J.L. & TECH. 12, 13, 35 (2015).

173. For a recent court discussion on disparate treatment, *see* Young v. United Parcel Serv., Inc., 135 S. Ct. 1338, 1345 (2015).

174. *Id*.

175. For a clear and recent explanation of the standards used to prove the different employment discrimination causes of action, *see generally* Deborah L. Brake, *The Shifting Sands of Employment Discrimination: From Unjustified Impact to Disparate Treatment in Pregnancy and Pay*, 105 GEO. L.J. 559 (2017).

176. Kim, *supra* note 106, at 866 (arguing that the disparate impact provision of Title VII should be interpreted and used to combat what she terms "classification bias," or employer reliance on "data algorithms, to sort or score workers in ways that worsen inequality or disadvantage along the lines of race, sex, or other protected characteristics.").

177. Barocas & Selbst, *supra* note 4, at 673–74.

An employer would clearly be engaging in disparate treatment if it used ATS reports as a mask to engage in forbidden discrimination.[178] Even if ATS reports scrub out references to protected characteristics, which is one of their selling points as a way to reduce employers' legal liability,[179] other variables in the report could serve as proxies for identifying group memberships. While this is certainly possible, employers are unlikely to use ATS as a cover for intentional discrimination. As Barocas and Selbst explain, "most cases of employment discrimination are already sufficiently difficult to prove; employers motivated by conscious prejudice would have little to gain by pursuing these complex and costly [data mining] mechanisms to further mask their intentions."[180] Even if an employer were purposefully discriminating, disparate treatment cases are hard to win, especially without a smoking gun, such as an employer's discriminatory comments. One would search in vain for a smoking gun in ATS reports, because "these models simply mine the available data, looking for statistical correlations that connect seemingly unrelated variables, such as patterns of social media behavior, with workplace performance."[181] Accordingly, ATS reports are more likely to reflect or foster implicit bias, through which unconscious stereotypes are used to make decisions. Although scholars have crafted compelling theories for finding implicit bias liability,[182] courts are very wary of recognizing this "second generation" form of discrimination. Instead, they are looking for bad actors with intentional animus.[183] Other data mining dangers are similarly devoid of conscious intent—such as erroneous data or replication of structural biases embedded within the workplace—and thus fail to meet the courts' demands.[184]

Because ATS use is unlikely to trigger conscious discrimination, disparate impact doctrine appears a better fit, but it is similarly constrained as a remedy. Under Title VII, an employer cannot use facially neutral policies or practices that have a disparate impact on a protected class, unless those policies or practices are justified by a legitimate business

178. Kim, *supra* note 106, at 884–85.

179. Barocas & Selbst, *supra* note 4, at 714.

180. Barocas & Selbst, *supra* note 4, at 693 (footnote omitted).

181. Kim, *supra* note 106, at 866.

182. *See generally e.g.*, Linda Hamilton Krieger & Susan T. Fiske, *Employment Discrimination Law: Implicit Bias and Disparate Treatment*, 94 CAL. L. REV. 997 (2006); Susan Sturm, *Second Generation Employment Discrimination: A Structural Approach*, 101 COLUM. L. REV. 458 (2001).

183. Brake, *supra* note 175, at 570–74.

184. Kim, *supra* note 106, at 887.

need that cannot be reasonably achieved by other means.[185] Even in the absence of big data, disparate impact cases are notoriously hard to win due to the complexity and expense of obtaining the necessary statistical evidence to demonstrate a disparate impact,[186] judicial biases against employment plaintiffs and a concomitant belief that discrimination is a relic of the past, a lack of understanding about how unconscious bias works,[187] and courts' willingness to accept employers' proffered justifications for business necessity.[188]

When big data is added to the mix, as Barocas and Selbst have pointed out, disparate impact is particularly difficult to establish.[189] This is because courts approve of employer hiring criteria that are job related, and computer models have the benefit of accessing massive amounts of data that are highly predictive of future performance.[190] Moreover, much data collection and mining incorporates unconscious biases that are baked into current structural disparities, making it hard for plaintiffs to identify alternative employment practices that achieve the same goals while being less discriminatory, as Title VII requires.[191]

---

185. This framework was established in *Griggs v. Duke Power Co*. 401 U.S. 424 (1971). The Supreme Court held that Title VII prohibits not only intentional discrimination, but also practices and policies that have a discriminatory impact upon employees. *Id*. The Court struck down an employer's job requirement for high school diplomas and certain test scores because there was an adverse impact on African-Americans and the requirements were not necessary to perform the power plant jobs at issue. *Id*.

186. *See* Sprague, *supra* note 172, at 40; Paul-Emile, *supra* note 141, at 926 ("[T]he fact that criminal records discrimination occurs almost exclusively during the hiring stage makes it difficult for an aggrieved applicant to acquire the empirical data necessary to show how the employer has treated similarly situated applicants.").

187. Disparate impact theory struggles to accommodate implicit bias claims because the legal standard focuses on a neutral employer policy, and thus fails "to consider the role of the individual decision-maker whose discrimination led to an adverse employment action." Audrey J. Lee, *Unconscious Bias Theory in Employment Discrimination Litigation*, 40 HARV. C.R.–C.L. L. REV. 481, 491 (2005).

188. *See* Michael Selmi, *Was the Disparate Impact Theory a Mistake?,* 53 UCLA L. REV. 701, 706 (2006); Michael Selmi, *Why are Employment Discrimination Cases So Hard to Win?*, 61 LA. L. REV. 555, 561–62 (2001); Ann C. McGinley, Ricci v. DeStefano*: Diluting Disparate Impact and Redefining Disparate Treatment*, 12 NEV. L.J. 626 (2001); Susan D. Carle, *A Social Movement History of Title VII Disparate Impact Analysis*, 63 FLA. L. REV. 251, 257 (2011); Barocas & Selbst, *supra* note 4, at 707 ("[C]ourts tend to accept most common business practices for which an employer has a plausible story.").

189. Barocas & Selbst, *supra* note 4, at 707–12 (describing the multiple hurdles for raising a disparate impact claim in the context of big data).

190. *Id*. at 707–08.

191. *Id*. at 710–11.

The challenges of establishing disparate impact claims are seen with regard to credit checks and criminal records—both of which may be part of ATS reports—but which also have longer histories as standalone pre-employment screening tools. Poor credit and criminal records are more prevalent in minority communities than among whites, thus their use has an adverse impact on the hiring of minorities.[192] Disparities in credit particularly impact African-Americans and result from historic patterns of discrimination that have created a racial wealth gap, as well as predatory lending practices targeting minority communities.[193] Similar disparities plague minorities with regard to criminal records, as a result of disproportionate policing and arrests in minority communities.[194] At the same time, studies do not find a predictive connection between these personal characteristics and job performance for most jobs.[195] Accordingly, civil rights groups have advocated against the use of both these criteria in hiring. [196]

In 2012, the Equal Employment Opportunity Commission (EEOC) issued guidance seeking to reduce the disparate impact of criminal background checks upon minorities[197] (the EEOC has not issued guidelines with regard to credit checks).[198] The criminal background check guidelines assert that it is unlawful for employers to adopt blanket no-hire policies for people with criminal backgrounds.[199] Instead, the EEOC advises employers to conduct individualized assessments of job candidates with criminal backgrounds prior to excluding them from a position. The

192. With regard to credit checks, *see* Pooja Shethji, Note, *Credit Checks Under Title VII: Learning from the Criminal Background Check Context*, 91 N.Y.U. L. REV. 989, 997–1002 (2016). Half of employers conduct credit checks in hiring. *Id*. at 991. With regard to criminal background checks, *see* Paul-Emile, *supra* note 141, 894–97; PERSIS S. YU & SHARON M. DIETRICH, NAT'L CONSUMER LAW CTR., BROKEN RECORDS: HOW ERRORS BY CRIMINAL BACKGROUND CHECKING COMPANIES HARM WORKERS AND BUSINESSES (2012), https://perma.cc/WQM7-TS83.

193. AMY TRAUB, DEMOS, DISCREDITED: HOW EMPLOYMENT CREDIT CHECKS KEEP QUALIFIED WORKERS OUT OF A JOB 7, 9 (2013), https://perma.cc/V9AD-G5L5.

194. *See* Paul-Emile, *supra* note 141, at 910–15.

195. *Id*. at 895 (criminal background checks); Shethji, *supra* note 192, at 991 (credit checks).

196. *See, e.g*., Lydell Bridgeford, *Q&A: Hiring Practices on the NAACP Legal Defense Fund's Radar*, BNA LAB. AND EMP. BLOG, Apr. 8, 2013, https://www.bna.com/qa-hiring-practices-b17179873189/. The "ban the box" movement advocates against required criminal background disclosures in job applications. *See* Dallan F. Flake, *When Any Sentence is a Life Sentence: Employment Discrimination Against Ex-Offenders*, 93 WASH. U. L. REV. 45, 92-95 (2015).

197. U.S. EQUAL EMP. OPPORTUNITY COMM'N, NO. 915.002 EEOC ENFORCEMENT GUIDANCE: CONSIDERATION OF ARREST AND CONVICTION RECORDS IN EMPLOYMENT DECISIONS UNDER TITLE VII OF THE CIVIL RIGHTS ACT OF 1964 (2012), https://perma.cc/Q4DN-DNZQ.

198. *See* Shethji, *supra* note 192, at 992.

199. *See supra* note 197.

guidance generated an immediate backlash from the business community, which complained about the "catch-22" of conducting criminal background checks and being held in violation of Title VII or failing to conduct a check and being held liable for negligent hiring.[200] Regardless of the merits of the guidance, subsequent EEOC disparate impact litigation based on credit and criminal background checks has faltered, as courts have found the plaintiff's proof inadequate.[201] At bottom, courts may be concerned about their role in remedying the complexities of societal discrimination that precede and extend far beyond an employer's decision to conduct a criminal background or credit check. Such reluctance is likely to plague evaluations of ATS systems, which aggregate multiple sources of data and thus attenuate the employer's responsibility for the patterns that emerge.

Social network data muddies the waters even more. Recently, some credit reporting companies have begun gleaning data from social media networks, thereby impacting the credit rating not only of the person subject to the report, but also the credit of third parties within that person's networks.[202] In other words, everyone's creditworthiness is impacting the entire network. Because ATS evaluations typically include credit information, it may be sweeping in these relational inferences from credit check companies, as well as generating them directly from social networks. This dynamic, in which social networks generate inferences about a person's job worthiness, raises the theoretical potential of associational discrimination claims under Title VII.[203] Under this theory,

---

200. HANS VON SPAKOVSKY, HERITAGE FOUND., THE DANGEROUS IMPACT OF BARRING CRIMINAL BACKGROUND CHECKS: CONGRESS NEEDS TO OVERRULE THE EEOC'S NEW EMPLOYMENT "GUIDELINES" (2012), https://perma.cc/T5UV-LJMG; John D. Bible, *To Check or Not to Check: New EEOC Enforcement Guidance on the Use of Criminal History Information in Making Hiring Decisions*, BUS. L. TODAY, MAR. 2013, https://www.americanbar.org/publications/blt/2013/03/05_bible.html.

201. *See, e.g.*, EEOC v. Kaplan Higher Educ. Corp., 748 F.3d 749 (6th Cir. 2014) (rejecting a disparate impact challenge to employer use of credit histories in hiring); EEOC v. Freeman, 961 F. Supp. 2d 783, 789 (D. Md. 2013), *aff'd*, 778 F.3d 463 (4th Cir. 2015) (rejecting a disparate impact challenge to employer use of criminal background checks).

202. Packin & Lev-Aretz, *supra* note 134, at 386.

203. While Title VII does not directly address this theory one way or the other, the Americans with Disabilities Act expressly forbids associational discrimination, meaning that an employee cannot be discriminated against for having an association with a person with a known disability. 42 U.S.C. § 12112(b)(4) (2012). The EEOC explains: "The purpose of the association provision is to prevent employers from taking adverse actions based on unfounded stereotypes and assumptions about individuals who associate with people who have disabilities." *Questions and Answers About the*

courts have recognized disparate treatment claims brought by plaintiffs whose employers took adverse actions against them because they had relationships with persons of another race or national origin.[204] For instance, courts have applied associational discrimination to cases in which a white man claimed he was not hired due to his marriage to a black woman,[205] a white woman asserted her employer refused to renew her contract because she associated with "Spanish citizens,"[206] and a white man alleged discriminatory discharge for having a biracial daughter.[207] As one court explained, "where an employee is subjected to adverse action because an employer disapproves of interracial association, the employee suffers discrimination because of the employee's *own* race."[208] These cases support the principle that employees should be able to associate with people of their own choosing without employers drawing negative conclusions from those relationships.

Big data discrimination may run afoul of this principle. Yet, despite its theoretical appeal, there are practical barriers to associational discrimination claims in the big data context. To begin with, successful associational discrimination cases involve intentional acts by employers motivated by conscious bias, and are thus framed as disparate treatment claims. By contrast, big data discrimination is likely to be unintentional, as data mining involves statistical correlations that do not require conscious efforts to target specific groups. Thus, it would be difficult, if not

---

Association Provision of the Americans with Disabilities Act, U.S. EQUAL EMP. OPPORTUNITY COMM'N (Feb. 2, 2011), https://perma.cc/Z8GY-M64K.

204. *See* Victoria Schwartz, *Title VII: A Shift from Sex to Relationships*, 35 HARV. J.L. & GENDER 209, 215–32 (2012) (summarizing the theory and cases); *see also* Matthew Clark, *Stating a Title VII Claim for Sexual Orientation Discrimination in the Workplace: The Legal Theories Available After* Rene v. MGM Grand Hotel, 51 UCLA L. REV. 313, 329 (2003).

205. Parr v. Woodmen of the World Life Ins. Co., 791 F.2d 888, 892 (11th Cir. 1986).

206. Reiter v. Ctr. Consol. School Dist., 618 F. Supp. 1458, 1459 (D. Colo. 1985).

207. *See* Tetro v. Elliot Popham Pontiac, Oldsmobile, Buick & GMC Trucks, Inc., 173 F.3d 988, 994 (6th Cir. 1999); s*ee also* Gresham v. Waffle House, Inc., 586 F. Supp. 1442, 1445 (N.D. Ga. 1984) (a white woman asserted she was fired due to her marriage to a black man); Barrett v. Whirlpool Corp., 543 F. Supp. 2d 812, 826 (M.D. Tenn. 2008) (a white woman claimed she suffered a hostile work environment due to her friendship with an African-American male co-worker); Deffenbaugh-Williams v. Wal-Mart Stores, Inc., 156 F.3d 581 (5th Cir. 1998), *vacated on other grounds by*, Williams v. Wal-Mart Stores, Inc., 182 F.3d 333 (5th Cir. 1999) (a white plaintiff claimed she was discharged due to her dating and marriage to a black person); Holcomb v. Iona Coll., 521 F.3d 130, 132 (2d Cir. 2008) (white basketball coach claimed discriminatory discharge based on his marriage to a black woman). The EEOC has applied the theory to sexual orientation discrimination as well. Baldwin v. Foxx, EEOC Appeal No. 0120133080, 2015 WL 4397641 (E.E.O.C. July 15, 2015), https://perma.cc/L8Q6-XHYC.

208. *Holcomb*, 521 F.3d at 139.

impossible, to ascertain the role that any specific social network data played in an ATS report that generated conclusions about a potential employee's likely job performance or tenure or other job-related characteristic. Job applicants would not explicitly be turned away from jobs because employers do not approve of their online friends. Rather, algorithms may conclude that the applicant's friends have characteristics that bear on the applicant's own suitability for the job. This is potentially problematic, but it does not fit within the current associational discrimination paradigm.

Discrimination law is also not a promising avenue when it comes to discrimination against the poor, whether intentional or unconscious. Employment law simply does not extend to discrimination on the basis of social class, even though decisions based on social media searches might "further disadvantage the poor by subjecting them to the negative judgments of those who control important resources[,]" such as employers.[209] For instance, obesity and smoking are more prevalent in low-income communities, and social media that reveals these traits may result in reputational harm that limits opportunities.[210] Certain clothing styles or social behaviors are also associated with poor communities and operate as signals of social class.[211] This is problematic because poverty carries social stigma in America—our governing ideology blames poverty on individual moral failings rather than structural dislocations in the economy.[212] Yet even intentional discrimination against the poor is perfectly legal and ATS will likely capture these indicia of poverty. Thus, these networked inferences, when implemented widely and without recourse, risk further dampening social mobility and trapping individuals in unemployment or low-wage employment.

B. Higher Education

1. Big Data Tools Impacting Access to Higher Education

Education is often cited as one of the primary pathways out of poverty, but increasingly, the online behavior of low-income applicants may

209. Thomas H. Koenig & Michael L. Rustad, *Digital Scarlet Letters: Social Media Stigmatization of the Poor and What Can Be Done*, 93 NEB. L. REV. 592, 611 (2015).

210. *Id*. at 600–01.

211. *Id*. at 595.

212. *Id*. at 598.

influence whether or not they are recruited for or ultimately gain entrance to college. Big-data driven insights are being used to help create efficiencies at various stages in the college admissions lifecycle, from marketing and recruitment to selection and retention. Some of these efforts are intended to reduce economic inequalities in access to higher education. For instance, in order to more effectively target college marketing materials to high-achieving, low-income students, certain institutions are using large-scale datasets to provide customized recommendations of "high-quality colleges and universities" for students who are likely to be admitted based on their previous academic performance.[213]

Other efforts are more focused on improving the admissions decision-making process and increasing graduation rates. As colleges increasingly look for ways to differentiate students, some are using predictive modeling tools that consider a wide range of factors beyond traditional application materials, such as how many friends and photos they have on social media platforms.[214]

The general practice of reviewing applicants' social media profiles is becoming more common among admissions officers. A recent Kaplan Test Prep survey of close to 400 college admissions officers across the United States found that 40% of admissions officers now say they visit applicants' social media pages to learn more about them, up from 10% in 2008.[215] Yet, the Kaplan questions measure a fairly basic form of checking up on students' digital footprints that still requires a certain level of judgment and discretion on the part of the reviewer. By contrast, when a student's social media data are fed into a third-party predictive analytics system, the reviewer may not understand what variables are factoring into a student's score or how each one is weighted and why.

This raises a number of ethical questions regarding the fairness, accuracy and transparency of this process: How can admissions staff be certain that a system has captured the correct social media profile for an applicant? How do these systems evaluate the value of a student's online network and interactions? Should a student be held liable for the way their extended family members or friends or other connections behave on social

213. Ben Castleman, *Big Data, Meet Behavioral Science*, BROOKINGS (Mar. 10, 2016), https://perma.cc/G2D2-QZQH.

214. Emmanuel Felton, *Colleges Shift to Using 'Big Data'—Including from Social Media—in Admissions Decisions*, HECHINGER REP. (Aug. 21, 2015), https://perma.cc/428S-QJLR.

215. Press Release, Kaplan, Kaplan Test Prep Survey: Percentage of College Admissions Officers Who Check Out Applicants' Social Media Profiles Hits New High; Triggers Include Special Talents, Competitive Sabotage (Jan. 13, 2016), https://perma.cc/TJ74-B3WR.

media? And how does the absence of a social media footprint (if one chooses to opt-out or restrict content) affect one's standing?

Even as these systems strive to create a better fit between prospective students and their schools, there can also be strong economic incentives driving the adoption of predictive modeling tools. The higher "yield" (i.e., demand) that a school can demonstrate to its creditors, the more it can save in reduced interest rates and improve overall rankings in lists such as the *U.S. News and World Report* ratings for universities.[216] IBM, one of the leading providers of big data-driven assessment tools for higher education, describes the social media value proposition as follows:

> Analyzing social media engagements not only provides insight on a candidate's personal interests, but, over time, analysis can also determine the behaviors of those who are likely to enroll and complete a degree program. Social media monitoring platforms generate real-time insights on content type and photos posted by current students and alumni, which can lead to a better idea of the kind of social media behavior to look for in a pool of candidates. In addition to peer interaction on their personal social media accounts, a candidate's interaction on a university's social media channels can be of interest to an admissions officer.[217]

It is unclear to what extent students realize that their social data streams might be fed into these kinds of assessments and how this awareness gap may affect different socioeconomic groups. Media reports have suggested that knowledge of social media monitoring by college admissions officers may be influencing more students to "clean up" their profiles before applying to college.[218] However, even if a student successfully sanitizes his or her own profile, the aforementioned method of social media monitoring may also include analysis of one's extended network of family and friends. Due to the networked nature of social media, that can create an extra layer of challenges to maintaining the kind of idealized online

---

216. Tim Lloyd, *How College Applications Change in the Era of Big Data*, MARKETPLACE (Jan. 14, 2014, 2:10 PM), https://perma.cc/C5P3-VMDV.

217. Lauren Willison, *3 Ways Universities Are Leveraging Big Data Analytics for Recruitment and Retention*, IBM BIG DATA & ANALYTICS HUB BLOG (Mar. 4, 2016), https://perma.cc/SC8U-5XUQ.

218. Natasha Singer, *Toning Down the Tweets Just in Case Colleges Pry*, N.Y. TIMES, Nov. 19, 2014, *available at* https://www.nytimes.com/2014/11/20/technology/college-applicants-sanitize-online-profiles-as-college-pry.html?_r=0.

presence that recruiters would like to see.

Former FTC Chairwoman Edith Ramirez has referred to this type of data-driven decision making as "data determinism," and cautioned that access to education may be one of the many areas impacted by a reliance on algorithmic assessments.

> Individuals may be judged not because of what they've done, or what they will do in the future, but because inferences or correlations drawn by algorithms suggest they may behave in ways that make them poor credit or insurance risks, unsuitable candidates for employment or admission to schools or other institutions, or unlikely to carry out certain functions.[219]

Scholars who study the role of social media in education have raised concerns about the fairness of reviewing personal profiles as part of the admissions process, and the ways in which it might discriminate against those who have a lower level of proficiency in social media privacy management.[220] As one dean of admissions, quoted in an article about big data-driven trends in college admissions, notes, "[t]his is the kind of stuff that savvy parents, students, and college counselors know about."[221] This begs the question: How are less savvy parents and students, and those who may not have regular engagement with college counselors, faring in this new environment? If lower-income social media users' profiles are more accessible than higher-income users, and analysis of their network connections is more likely to reflect an entrenched set of structural disadvantages, then the conclusions drawn from these analyses will replicate those inequities.

2. Legal Analysis of Predictive Analytics in College Admissions

The use of predictive analytics generated from big data sources such as social media postings, test scores, and demographic data faces few legal limits. No law prohibits colleges from gathering information about students from social media or other publicly available information.

With regard to the use of data, there are two main legal doctrines

---

219. Edith Ramirez, Chairwoman, Fed. Trade Comm'n, Keynote Address at the Technology Policy Institute Aspen Forum: Privacy Challenges in the Era of Big Data: A View from the Lifeguard's Chair (Aug. 19, 2013), https://perma.cc/5VA9-8U97.

220. Rey Junco, *The Ethics of Facebook-Stalking University Applicants*, SOC. MEDIA HIGHER EDUC. (Nov. 8, 2012), https://perma.cc/UJM6-5ET7.

221. Felton, *supra* note 214.

potentially at issue: Title VI of the Civil Rights Act,[222] which prohibits discrimination in education, and the Family Education Rights and Privacy Act (FERPA),[223] which governs the confidentiality of student records.

One major concern about predictive analytics in the higher education sphere is that minority college applicants will be disproportionately excluded from admissions. Colleges are looking to identify metrics of student retention and success,[224] and the metrics they are using may unintentionally harm minorities. This is a problem of disparate impact, in which a facially neutral process has a differential impact on minority groups.[225] The gathering and analysis of big data can look particularly "neutral," given that computers do the work based on seemingly objective criteria. This neutrality may be a mirage, however, because software engineers craft code that can unintentionally embed social and cognitive biases into the analytics. If minorities are denied admission as a result, they might have a disparate impact claim under Title VI of the Civil Rights Act (similar to the sort of claim discussed above in connection with employment).

Title VI provides that "[n]o person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance."[226] It covers almost all colleges and universities, both public and private.[227] Title VI allows individuals to seek relief in court for intentional discrimination, but it does not include language barring disparate impact.[228] Nevertheless, the Department of Education (DOE) has adopted disparate impact theory by issuing regulations that prohibit practices that have the "effect of" discriminating on the basis of race, color, or ethnicity.[229]

Only the DOE can enforce a Title VI disparate impact claim; private

222. 42 U.S.C. § 2000d (2012).

223. 20 U.S.C. § 1232g (2012).

224. *See* Ry Rivard, *Predicting Where Students Go*, INSIDE HIGHER ED (Sept. 19, 2014), https://perma.cc/F5RR-Z5UB.

225. *See supra* note 177 (describing disparate impact in the employment context).

226. 42 U.S.C. § 2000d (2012).

227. Office for Civil Rights, *Race and Origin Discrimination: Frequently Asked Questions*, U.S. DEP'T OF EDUC., https://perma.cc/YZ3Q-5B86 (last visited June 22, 2017).

228. Olatunde C.A. Johnson, *The Agency Roots of Disparate Impact*, 49 HARV. C.R.-C.L. L. REV. 125, 130 (2014) (discussing statutory and regulatory interpretations of Title VI).

229. 34 C.F.R. § 100.3(b)(2).

citizens cannot.[230] A student or other individual can trigger a DOE investigation by filing a complaint with the DOE's Office of Civil Rights.[231] The test for disparate impact claims in the educational context is borrowed from the employment context of Title VII.[232] Under this test, the plaintiff must first show that the practice in question results in a significant disparity in the provision of a benefit or service that is based on race or national origin.[233] The burden then shifts to the college, who must demonstrate that the policy has a "substantial legitimate justification."[234] If the college meets its burden, it can still be liable if alternative practices exist that would meet the college's educational goals and result in lower disparities.[235]

Using this framework, scholars and advocates have crafted disparate impact theories designed to challenge the use of SAT scores as well as criminal background checks in admissions, both of which may disproportionately exclude minorities. In brief, the argument is that use of these data points have a disproportionately harmful impact on black and Hispanic students, that the data is correlated to low-income status but not to future success (for SATs) or unsafe behavior (for criminal background), and that there are less discriminatory alternatives for predicting educational outcomes (the college's justification for SAT scores) and student safety (the justification for criminal background checks).[236]

In a legal challenge to the use of the SATs in the Title IX context (prohibiting educational discrimination on the basis of gender), a federal court ruled that New York State's exclusive reliance on SAT scores to

230. Alexander v. Sandoval, 532 U.S. 275, 281 (2001) (holding that the disparate impact Title VI regulations are not privately enforceable). For a critique of *Sandoval*, *see* Johnson, *supra* note 228, at 131–32.

231. Office for Civil Rights, *Education and Title VI*, U.S. DEP'T OF EDUC., https://perma.cc/JH6N-XXX9 (last visited June 22, 2017).

232. *See* Rebecca R. Ramaswamy, *Bars to Education: The Use of Criminal History Information in College Admissions*, 5 COLUM. J. RACE & L. 145, 154–55 (2015) (describing the burden shifting framework for disparate impact set forth in Wards Cove Packing Co. v. Antonio, 490 U.S. 642 (1989)); U.S. DEP'T OF JUSTICE, TITLE VI LEGAL MANUAL 49–50 (2001).

233. *See, e.g.*, Powell v. Ridge, 189 F.3d 387, 393 (3d Cir. 1999) ("[T]he courts of appeals have generally agreed that the parties' respective burdens in a Title VI disparate impact case should follow those developed in Title VII cases."); Georgia State Conference of Branches of NAACP v. Georgia, 775 F.2d 1403, 1418 (11th Cir. 1985) (applying Title VII's burden-shifting test to Title VI disparate impact litigation); Larry P. v. Riles, 793 F.2d 969, 982 n.9 (9th Cir. 1984).

234. Powell, 189 F.3d. at 393.

235. *Id*. at 394.

236. *See* Ramaswamy, *supra* note 232, at 154–62 (regarding criminal background checks); Kimberly West-Faulcon, *The River Runs Dry: When Title VI Trumps State Anti-Affirmative Action Laws*, 157 U. PA. L. REV. 1075, 1124–1131 (2009) (regarding SAT scores).

award merit-based college scholarships had a disparate impact on women.[237] Similarly, the DOE's Office of Civil Rights is currently investigating a Florida scholarship program that relies heavily on strict SAT and ACT score cut-offs as criteria and that awards most scholarships to white and affluent families.[238] While no legal case has been brought testing this theory against criminal background checks, there is sustained advocacy in this area. In February 2016, the Lawyers' Committee for Civil Rights Under Law called on the organization that administers the Common Application—a standardized college application form used by more than 600 colleges—to cease using questions about applicants' educational disciplinary histories, criminal records, and juvenile justice backgrounds.[239] In 2015, New York University announced that it would review applications "without awareness of whether the applicant checked the box" regarding past crimes, and if a student with a criminal background passes that stage, the decision then goes to a team of admissions officers "specially trained on fact-based assessment and issues of bias."[240] The University explained that this would strike a balance between educating a diverse group of students and ensuring the safety of the campus community.[241]

Ultimately, grassroots advocacy might be the best strategy for challenging big data in admissions. As noted above with regard to employment, disparate impact cases are generally hard to win.[242] To begin with, there is no protection for discrimination based on poverty, even though it is disproportionately associated with minorities. Even if any adverse impact of big data in admissions falls clearly on minorities (as opposed to poor applicants in general), a plaintiff who suspects he or she is a victim of disparate impact through predictive analytics would have to convince DOE to investigate the case. Even if DOE pursued a case, litigation in this area is difficult to win due to the statistical demands of proving disparate impact, as well as judicial deference to college

237. Sharif v. New York State Educ. Dep't, 709 F. Supp. 345 (S.D.N.Y. 1989).

238. David Smiley, Michael Vasquez & Kathleen McGrory, *Feds Investigate Florida's Bright Futures Scholarships*, MIAMI HERALD (Mar. 22, 2014, 4:00 PM), https://perma.cc/D2SQ-FB5Z.

239. Press Release, Stacie Burgess, Lawyer's Comm. for Civil Rights Under Law, Lawyers Committee Calls for the Common Application to Eliminate Discriminatory Barriers to College Admissions Nationwide (Feb. 18, 2016), https://perma.cc/FS5W-RCFS.

240. Press Release, N.Y. Univ., NYU Revises Admissions Practices for Applicants Convicted of a Crime (May 23, 2015), https://perma.cc/969W-U5FG.

241. *Id*.

242. *See supra* notes 186–188 and accompanying text.

determinations about educational necessity.[243] Further, the very theory of disparate impact in the educational context is highly contested given that the statute does not explicitly permit it. Big data makes proof of disparate impact even harder. The process of aggregation makes it difficult to identify a specific variable that is linked to a disparate impact, as courts require.[244] As Jules Polonetsky and Omer Tene write, "the hallmark of big data is the escape of information from the confines of a structured database and the ability to harvest, analyze, rearrange, and reuse freestanding information."[245]

Not surprisingly then, the legal and social movements against the use of SATs and criminal background checks do not tackle the far murkier issue of how social media information is being used to make decisions about student admissions. Pointing to the SATs as a discriminatory culprit is straightforward; disentangling a predictive score generated from big data is more complex. Ironically, tests like the SATs were initially created to provide "neutral" metrics of merit in order to expand educational opportunities to a broader range of students,[246] yet the reality is that test results reflect the disadvantages of test takers.[247] The use of big data is likely to mirror this trajectory. Lower income young adults are avid social media users, but less likely than their higher income peers to use privacy settings to limit the information they share online.[248] Will SAT scores and criminal background checks be replaced by social media and data broker proxies?

It is possible that a student could uncover the variables that influenced his or her admissions decision using FERPA,[249] which protects the confidentiality of student educational records. Schools that violate FERPA risk a loss of federal funding, although this penalty has never been

243. *See* Daniel Kiel, *No Caste Here? Toward a Structural Critique of American Education*, 119 PENN. ST. L. REV. 611, 631 (2015).

244. *See* Kim, *supra* note 106, at 51.

245. Jules Polonetsky & Omer Tene, *Who is Reading Whom Now: Privacy in Education from Books to MOOCs*, 17 VAND. J. ENT. & TECH. L. 927, 962 (2015).

246. West-Faulcon, *supra* note 236, at 1113–14.

247. *See* Dimitrios Halikias & Richard V. Reeves, *Race Gaps in SAT Scores Highlight Inequality and Hinder Upward Mobility*, BROOKINGS (Feb. 1, 2017), https://perma.cc/X6XK-AHMZ.

248. Results from the main survey discussed in this paper indicate that, among young adult Internet users ages eighteen to twenty-nine living in households earning less than $40,000 per year, 73% have used privacy settings to limit who can see what they post online. By comparison, fully 85% of young adult Internet users living in households earning $40,000 or more per year said they use privacy settings. Mary Madden, *supra* note 77. They also receive less education about privacy settings from teachers and parents than their wealthier peers. *See* Koenig & Rustad, *supra* note 209, at 611–12.

249. 20 U.S.C. § 1232g (2012); 34 C.F.R. § 99.1.

imposed.[250] Under FERPA, students have the right to review educational records, request corrections, receive a hearing if the correction is denied, prohibit the release of personally identifiable information, and receive an annual notice of FERPA rights.[251] However, even if a student gained access to their file, he or she would be hard-pressed to identify an underlying algorithmic variable or how it impacted the school's decision.

As with many digital records, there are concerns that the information feeding predictive analytics is not always accurate. The law provides little recourse. FCRA does not apply to admissions processes at educational institutions.[252] Moreover, many colleges gather student data through student self-reporting or by searching for information themselves,[253] thus taking them outside FCRA's regulation of credit reporting. If colleges were subject to FCRA with regard to prospective students, they would have to notify applicants about a background check and obtain their consent, as well as notify a student before rejecting them if the decision were based in part on the report, and provide the student with the opportunity to challenge the accuracy of the report.[254] Such changes would be a positive step for ensuring fairness in higher education admissions, but could face many of the limitations of FCRA in the consumer context.

One trend in higher education that is on the FTC's radar is the practice of "lead generation," which may prey upon low-income students—particularly those whose social media activity is readily accessible. As the FTC defines it, lead generation "is the practice of identifying or cultivating consumer interest in a product or service, and distributing this information to third parties."[255] The FTC notes that such leads often contain sensitive personal and financial information "that may travel through multiple online marketing entities before connecting with the desired business."[256]

---

250. *See* Polonetsky & Tene, *supra* note 245, at 967.

251. 20 U.S.C.S. § 1232g(a). Parents hold these rights for their children until they are eighteen or enroll in college, at which time they transfer to the student. *Id*. Polontesky & Tene point out that privacy of student records may have the cost of keeping important data about structural disadvantage out of the hands of civil rights organizations and educational reformers. Polonetsky & Tene, *supra* note 245, at 969.

252. *See* Darby Dickerson, *Background Checks in the University Admissions Process: An Overview of Legal and Policy Considerations*, 34 J.C. & U.L. 419, 460 (2008).

253. *See supra* notes 219 and 224 and accompanying text.

254. *See supra* notes 135–138 and accompanying text.

255. *Follow the Lead: An FTC Workshop on Lead Generation*, FED. TRADE COMM'N (Oct. 30, 2015), https://perma.cc/7WP3-CZTJ.

256. *Id*.

In the higher education sphere, there are websites that gather information about potential students and then sell the data to for-profit colleges, where students often assume crippling debt with few job prospects after graduation.[257] The for-profit educational industry, in particular, targets low-income students.[258] In April 2016, the FTC announced a settlement of its first enforcement action against a lead generator in the educational sphere.[259] The FTC alleged that Gigats.com had deceived consumers into thinking they were being pre-screened for jobs when it was instead gathering information for for-profit educational schools that were paying for the leads.[260] While transparent and truthful lead generation has the potential to connect low-income students to educational opportunities, continued FTC enforcement in this area will be important to protect students from deceptive practices. At the same time, lead generation seems to be only the tip of the iceberg in terms of potential big data disparities in the field of higher education.

C. Policing

1. The Emerging World of Threat Scores and Predictive Policing Tools

Due to historic patterns in law enforcement, the problems that algorithmic decision-making (drawing on social media and Internet use) poses for the poor are especially acute when considering the potential consequences of "threat scoring" systems and other predictive policing tools. As noted in the May 2014 White House Report *Big Data: Seizing Opportunities, Preserving Values*, one of the more controversial features of new predictive policing tools is the ability to create individualized scores to assess a single person's propensity for being involved in a crime.[261] While the formulas behind proprietary analytical models are currently inaccessible to researchers or even the police departments who purchase these tools, some insights can be gleaned through analysis of

257. CTR. FOR DIG. DEMOCRACY & U.S. PIRG EDUC. FUND, PRIVATE FOR-PROFIT COLLEGES AND ONLINE LEAD GENERATION: PRIVATE UNIVERSITIES USE DIGITAL MARKETING TO TARGET PROSPECTS, INCLUDING VETERANS, VIA THE INTERNET (2015), https://perma.cc/38U4-CPN5.

258. *See* Alia Wong, *The Downfall of For-Profit Colleges*, ATLANTIC (Feb. 23, 2015), https://perma.cc/7N5A-JGX5.

259. Press Release, Fed. Trade Comm'n, FTC Charges Education Lead Generator with Tricking Job Seekers by Claiming to Represent Hiring Employers (Apr. 28, 2016), https://perma.cc/3UVS-PK8D.

260. *Id*.

261. SEIZING OPPORTUNITIES, *supra* note 57, at 31.

marketing materials associated with some of the platforms being sold for use in law enforcement across the United States.

In a recent review of promotional materials and public statements associated with *Beware*, one of the controversial threat scoring tools being piloted in places such as Fresno, California, David Robinson noted that, "threat scores may reflect everything from criminal histories to social media activity to health-related history."[262] Rather than relying on police department or city records, the assessments are based on data gathered from commercial data brokers.[263] However, it is entirely unclear—both to police departments and the general public—how individual variables being fed into these models might be weighted and what kind of threat they are actually measuring.[264] Those who are subject to these analyses (which are run in response to a 911 call) are given neither any insight into how the data is being used, nor any ability to correct errors in cases where inaccurate information may result in a mistaken assessment of the individual.[265]

With many different data streams feeding into these models, there are myriad opportunities for problems associated with information quality and accuracy. In one widely reported anecdote, Fresno, California councilman Clinton J. Olivier asked to have his threat score run by the Beware system at a local hearing last November, the tool returned a threat level of "yellow" for his address (rather than "green," as one might expect of a public official)—possibly due to the activities of someone who previously lived at his address.[266]

Problems with information quality that can lead law enforcement to make inaccurate assumptions about the risks associated with a given address are not unique to big data analytics, but can be exacerbated by the transient nature of low-income communities. A recent *Washington Post*

---

262. David Robinson, *Buyer Beware: A Hard Look at Police 'Threat Scores*,' EQUAL FUTURE (Jan. 14, 2016), https://perma.cc/4MSE-JVXN.

263. *Id*.

264. This lack of transparency was highlighted as a major concern by the ACLU last year, who, along with a coalition of civil rights groups and technology companies, published a joint statement to express their shared concerns about predictive policing tools. *See* THE LEADERSHIP CONFERENCE ON CIVIL & HUMAN RIGHTS ET AL., PREDICTIVE POLICING TODAY: A SHARED STATEMENT OF CIVIL RIGHTS CONCERNS (2016), https://perma.cc/8UPU-J783.

265. ALEXANDRA MATEESCU ET AL., DATA & SOC'Y RESEARCH INST., SOCIAL MEDIA SURVEILLANCE AND LAW ENFORCEMENT (2015), https://perma.cc/S4Z7-634X.

266. Justin Jouvenal, *The New Way Police Are Surveilling You: Calculating Your Threat 'Score,'* WASH. POST (Jan. 10, 2016), https://perma.cc/GQH4-BXJV.

article relayed the story of a 63-year-old grandmother, Sallie Taylor, whose home was wrongfully raided by D.C. police one evening in January 2015. After bursting through her door, pointing a shotgun at her face and pushing her to the floor, the officers searched through her belongings for thirty minutes and ultimately left empty-handed. Police later discovered that, due to outdated information in the court records system, they had used the wrong address for the suspect they had been searching to find.[267] Similar scenarios occurred eleven more times in Washington, D.C. over a two-year period.[268]

While threat scoring generally targets a specific residence or person, other predictive policing tools offer a broader portrait of potential crime "hot spot" locations that are intended to help police better allocate their attention and resources.[269] Early versions of these systems required a police investigator's input into the weighting of the variables and decisions about how they would be factored into the model, but newer systems increasingly rely on machine learning and "[don't] require a human to figure out what variables matter and how much."[270]

One such example is the Hitachi Visualization Predictive Crime Analytics tool, which "gobbles massive amounts of data—from public transit maps, social media conversations, weather reports, and more—and uses machine learning to find patterns that humans can't pick out."[271] In particular, the role of social media input has been described as an especially important component of the tool, increasing its predictive accuracy by 15%.[272]

Location data gleaned from social media posts can be especially valuable in big data-driven policing tools. As Mateescu et al. note in their discussion of social media surveillance tools currently being used by law enforcement,

> Companies like Geofeedia offer products that use the location data

267. John Sullivan, Derek Hawkins & Pietro Lombardi, *Probable Cause*, WASH. POST (Mar. 5, 2016), https://perma.cc/2GBQ-8SAX.

268. After a review of 2000 warrants served by D.C. police between January 2013 and January 2015, the *Post* found that there were a dozen misdirected raids where "officers acted on incorrect or outdated address information." *Id*.

269. SARAH BRAYNE ET AL., DATA & SOC'Y RESEARCH INST., PREDICTIVE POLICING (2015), https://perma.cc/WC5L-88RS.

270. Sean Captain, *Hitachi Says It Can Predict Crimes Before They Happen*, FAST COMPANY (Sept. 28, 2015, 9:00 AM), https://perma.cc/MVY3-DGHF.

271. Amy X. Wang, *Hitachi Says It Can Predict Crimes Before They Happen*, QUARTZ (Sept. 29, 2015), https://perma.cc/DJ7D-U5H6.

272. *See supra* note. 270.

> of social media posts, when available, and map them. Using these maps, clients are able to specify a delimited geographic area and view all geotagged posts coming from that location in near real-time. Use of geotagging features to map social media activity has been touted as a crucial tool in assisting first responders in emergencies, as well as surveilling areas of concentrated activity, such as concerts or public protests. [273]

However, the availability of social media as a viable input depends on both a high level of usage among the communities under surveillance and the use of public platforms or public settings to make the communications broadly accessible to law enforcement. While police departments routinely request social media data from specific accounts as part of ongoing investigations, they are generally not able to monitor the complete firehose of social media data (both public and private) in real-time. Typically, social media monitoring systems rely on *public* social media data streams and this allows them to avoid what is seen as a "legal gray area" around content not intended for public consumption.[274] Given the aforementioned propensity of lower-income social media users to post content publicly and their tendency to rely on mobile devices that may be more vulnerable to law enforcement surveillance, the effective impact of this monitoring is not likely to be evenly distributed.

NC4, which produces the NC4 Signal tool to provide social media monitoring for law enforcement, is also designed to "listen" to public communications across popular social media platforms, such as Twitter and Facebook. The resulting cache of data gathered through the system includes text, images, and video that can be filtered and visualized to enhance "operational decision-making to ensure optimum results."[275] A blog post on the company's website lists the "pros and cons" of social media monitoring and touts the operational benefits of the method, but concedes that it can raise some privacy and perception-related concerns: "there aren't any true operational drawbacks to using social media monitoring software, but it can create problems with reputation and perception in the community, as well as with privacy advocates."[276] One

273. Mateescu, *supra* note 265, at 4 (footnotes omitted).

274. *See, e.g.*, *A Look at the Pros and Cons of Social Media Monitoring in Law Enforcement*, NC4, https://perma.cc/P72X-DWGY.

275. *Leveraging Open Source Intelligence (OSINT) for Public Safety,* NC4, https://perma.cc/6QCA-XX6J (last visited Jun. 22, 2017).

276. *See supra* note 274.

such problem is the risk of misinterpretation; when social media posts are analyzed outside of the original context in which they were shared, jokes, memes, sarcasm, and irony are not taken into account. In addition, individuals may be assessed not only by what they choose to post, but also by the content that their friends or followers post in association with their profile. These problems can be heightened when determinations are made largely by algorithms and automated systems, which can exacerbate entrenched biases. As Mateescu et al. argue, the removal of human judgment in the assessment of risk creates the potential for more widespread systematic bias.

Absent careful review, machine learning techniques applied to social media could easily reinforce existing patterns of enforcement, which partly reflect a disproportionate focus on people of color. To the extent that they replace human discretion, these automated systems may be trading individual bias–malicious or otherwise–for a new, systematic bias.[277]

2. Legal Analysis of Predictive Policing

Existing law fails to protect citizens from the potential dangers of predictive policing. Simply put, "[t]he rise of big data technologies offers a challenge to the traditional paradigm of Fourth Amendment law."[278] The Fourth Amendment protects citizens from unreasonable government searches and seizures.[279] The Supreme Court has long ruled that the Fourth Amendment protects only objectively reasonable expectations of privacy. In *Katz v. United States*, the Court held that a defendant had a reasonable expectation of privacy in a phone booth.[280] Once outside the proverbial phone booth, a citizen loses this protection. Thus, there is no Fourth Amendment protection for information people share in public or to third parties,[281] such as "data given to commercial third parties, including banking records, telephone call lists, cell phone locations, or Internet

277. Mateescu, *supra* note 265, at 7.

278. Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 329 (2015).

279. U.S. CONST. amend. IV. ("The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.").

280. 389 U.S. 347, 359 (1967).

281. United States v. Miller, 425 U.S. 435, 443 (1976) ("The depositor takes the risk, in revealing his affairs to another, that the information will be conveyed by that person to the Government.").

search or subscriber information." [282]

Scholars have argued that this doctrine "should not hold its traditional force once the police deploy the tools of big data,"[283] and the Supreme Court has signaled that the law will need to adapt to emerging technologies. In *United States v. Jones*, the Court held that police placement of a GPS device on a suspect's car without a warrant violated the Fourth Amendment because it was a physical intrusion that violated the defendant's right to privacy.[284] In this narrow ruling, the Court did not endorse the mosaic theory, which the appellate court below had applied.[285] The mosaic theory provides that prolonged surveillance generates individual pieces of information that constitute a Fourth Amendment search when aggregated.[286] In other words, the "whole reveals far more than the individual movements it comprises."[287] While the Supreme Court majority did not adopt this theory, five Justices intimated they are poised to reconsider the notion of reasonable expectations in our digital age. As Justice Sotomayor stated in concurrence, "I would ask whether people reasonably expect that their movements will be recorded and aggregated in a manner that enables the Government to ascertain, more or less at will, their political and religious beliefs, sexual habits, and so on."[288] At this time, however, the third-party doctrine prevails.

Statutes are no bulwark against predictive policing given that existing legal protections for health and financial information typically yield to law enforcement needs.[289] Moreover, if government obtains its information from a private data broker such as Beware, even the existing minimal statutory requirements for legal process can be evaded.[290] And of course, the Constitution does not constrain the behavior of private data brokers, only government actors.

Once the police have data in their possession, the issue becomes how they can use it. Police may initially want to use predictive analysis to conduct surveillance of "persons suspected of ongoing or future criminal

---

282. *See* Ferguson, *supra* note 278, at 373–74 (footnotes omitted); *see also* Elizabeth E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 WASH. L. REV. 35, 59 (2014).

283. *Joh, supra* note 282, at 62 (footnote omitted).

284. 565 U.S. 400, 404 (2012).

285. *See id*.; United States v. Maynard, 615 F.3d 544, 562 (D.C. Cir. 2010).

286. *Maynard*, 615 F.3d at 562.

287. *Id*.

288. *Jones*, 565 U.S. at 416 (Sotomayor, J., concurring).

289. *See* Ferguson, *supra* note 278, at 374.

290. *Id*. at 360, 379.

activity."[291] This surveillance phase of policing, undertaken prior to a search, detention, or arrest, is not a Fourth Amendment event, and it thus does not require individualized suspicion.[292] As Elizabeth Joh has explained, police have the unfettered discretion to single out certain people for investigation based on predictive tools.[293] Surveillance could potentially bump up against other constitutional rights, but relief through these avenues is unlikely. Claims that discretionary surveillance violates First Amendment rights of association, such as with police surveillance of a group engaging in political activity, often falter on standing grounds.[294]

Following surveillance, police may want to stop, question, search, or seize a suspect. These phases of police activity require individualized suspicion. For a stop and frisk, the police must have a reasonable suspicion, and for a more sustained search or an arrest, police must have probable cause (and sometimes a warrant). Courts look at a totality of the circumstances to determine whether these standards of individualized suspicion have been met.[295] Traditionally, police developed reasonable suspicion through observing activities in the real world and gathering information from sources. The process was retrospective and particularized, whereas predictive policing is generalized and prospective.[296]

As Andrew Ferguson explains, the use of big data in predictive policing "has the potential to change the reasonable suspicion calculus because more personal or predictive information about a suspect will make it easier for police to justify stopping a suspect."[297] Big data not only adds factors to the totality of circumstances test, but it can also be particularized to a specific subject, as courts demand.[298] Yet increased granularity in data can increase the risks of disparate impact due to inadequacies and inaccuracies in data.[299] Ferguson thus argues that courts should require a

291. *See* Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL'Y. REV. 15, 18 (2016).

292. *Id*. at 17.

293. *Id*. at 34.

294. *Id*. at 35.

295. *See* Ferguson, *supra* note 278, at 339.

296. *See* Fabio Arcila, Jr., *Nuance, Technology, and the Fourth Amendment: A Response to* Predictive Policing and Reasonable Suspicion, 63 EMORY L.J. ONLINE, 2087, 2090 (2014).

297. Ferguson, *supra* note 278, at 351. *See also* Arcila, *supra* note 296, at 2090; Joh, *supra* note 282, at 28.

298. Ferguson, *supra* note 278, at 387. *See also* Andrew D. Selbst, *Disparate Impact in Big Data Policing*, GA. L. REV. 38 (forthcoming 2017) ("[I]f a model is individualized enough, it satisfies the Fourth Amendment.").

299. Selbst, *supra* note 298, at 39–40.

direct link between predictive data about a suspect and police suspicion arising from direct observation.[300] Whether courts will heed this advice remains to be seen.

Equal Protection claims based on discriminatory police enforcement resulting from big data would also likely fail.[301] Predictive analytics might embed stereotyped views of suspects, but it does so unintentionally.[302] The Court has held that discriminatory intent is required to prevail under the Equal Protection Clause.[303] Indeed, even old-fashioned exercises of police discretion that result in disparate impacts on minorities are nearly impossible to challenge under equal protection.[304]

Another concern regarding predictive analytics is the large number of mistakes within the data.[305] In fact, the FBI's files, which are regularly used for background checks, are known to contain hundreds of thousands of mistakes.[306] In addition, programmers can make mistakes in creating the software and algorithms that analyze the data.[307] However, there are no meaningful quality controls on shared data and no individual rights to learn about or correct mistakes.[308] As a constitutional matter, the Supreme Court holds that exclusion of evidence from trial that was based on erroneous evidence in government databases is required only in cases of gross negligence or systemic misconduct.[309] Meeting this standard is nearly impossible given the difficulty of isolating erroneous material from the reams of data involved, as well as the proprietary nature of and secrecy underlying the databases.[310] In short, whether big data is accurate or error-

300. Ferguson, *supra* note 278, at 388.

301. *See* Selbst, *supra* note 298, at 5, 29.

302. *Id*. at 29.

303. Washington v. Davis, 426 U.S. 229 (1976). For a prominent critique, see Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987).

304. *See* Paul Butler, *The White Fourth Amendment*, 32 TEX. TECH. L. REV. 245, 246 (2010); Kevin R. Johnson, *How Racial Profiling in America Became the Law of the Land:* United States v. Brignoni-Ponce *and* Whren v. United States *and the Need for Truly Rebellious Lawyering*, 98 GEO. L.J. 1005, 1006 (2010) ("[R]acial profiling by law enforcement authorities in the United States has long been permitted and encouraged, if not expressly authorized, by U.S. constitutional law.").

305. Ferguson, *supra* note 278, at 398.

306. *Id*. at 399.

307. *See* Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 925 (2016).

308. Ferguson, *supra* note 278, at 398–99.

309. *See* Rich, *supra* note 307, at 925–27 (citing Arizona v. Evans, 514 U.S. 1, 15–16 (1995); Herring v. United States, 555 U.S. 135, 146 (2009)).

310. *Id*. at 928.

filled, current law provides scant constraint on its use in the criminal justice system.

Predictive policing's use of social media information gleaned from on-line networks also raises the specter of guilt by association, which is disfavored in our justice system. As the Supreme Court has stated, guilt by association is "alien to the traditions of a free society."[311] At the same time, "it is beyond debate" that the right of free association "for the advancement of beliefs and ideas is an inseparable aspect of . . . `liberty.'"[312] However, the right of free association, as currently conceived, is largely focused on protecting the ability of formally constituted groups to function without government interference.[313] For instance, foundational Supreme Court cases protected the NAACP from having to turn over its membership lists to the government.[314] By contrast, many social media relationships are between loose networks of friends and acquaintances and do not fit this paradigm.[315] Furthermore, free association doctrine protects communications that are either expressive (meaning political in nature) or intimate (such as within families),[316] but not social ones that constitute the bulk of web-based chatter.[317] In addition, algorithms scoop in metadata and other non-content data, which are not words and thus do not have an associative dimension as understood in traditional doctrine.[318] For all these reasons, free association doctrine will need considerable revamping to protect against guilt by on-line association.[319]

---

311. NAACP v. Claiborne Hardware, 458 U.S. 886, 932 (1982).

312. NAACP v. Alabama ex rel. Patterson, 357 U.S. 449, 460 (1958).

313. Boy Scouts of America v. Dale, 530 U.S. 648, 655–56 (2000).

314. *NAACP v. Alabama ex rel. Patterson*, 357 U.S. at 466.

315. Peter Swire writes about the tension between protecting individual privacy (which calls for data protection) and encouraging online political activity (which calls for data empowerment). Peter Swire, *Social Networks, Privacy, and Freedom of Association: Data Protection vs. Data Empowerment*, 90 N.C. L. REV. 1371, 1376–77 (2012). This is a different issue than guilt by algorithm in that the communications Swire is concerned with are political in purpose and generally not leading to punitive measures. *Id*. at 1377–80.

316. Roberts v. U.S. Jaycees, 468 U.S. 609, 618 (1984); John D. Inazu, *Virtual Assembly*, 98 CORNELL L. REV. 1093, 1099–1100 (2013).

317. *See* Inazu, *supra* note 316, at 1119; Deven R. Desai, *Constitutional Limits on Surveillance: Associational Freedom in the Age of Data Hoarding*, 90 NOTRE DAME L. REV. 579, 593–94 (2014).

318. *See* Desai, *supra* note 317, at 589; Katherine J. Strandburg, *Freedom of Association in a Networked World: First Amendment Regulation of Relational Surveillance*, 49 B.C. L. REV. 741, 747 (2008).

319. For scholars attempting to expand freedom of association to cover on-line communications, *see, e.g.*, Katherine J. Strandburg, *Membership Lists, Metadata, and Freedom of Association's Specificity Requirement*, 10 I/S: J.L. & POL'Y FOR INFO. SOC'Y 327 (2014); Desai, *supra* note 317; Chris J. Chasin, *The Revolution Will Be Tweeted, But the Tweets Will Be Subpoenaed: Reimagining*

IV. SUGGESTED REMEDIES AND THEIR EFFICACY FOR LOW-INCOME POPULATIONS

Current legal frameworks offer little protection or recourse for privacy-related harms experienced by the poor. At present, privacy law in the United States is fragmented and sectoral.[320] Unlike most other developed nations, the United States does not have a single data protection law.[321] Instead, we have industry-specific statutory protections, such as laws that govern the health or financial services industries.[322] Outside of these narrow statutes, the United States relies primarily on self-regulation by the entities that gather and maintain personal data and puts the onus on individuals to police their own data disclosures.[323] Given the gaps in United States privacy law, a diverse group of scholars, policymakers, and practitioners have made numerous recommendations to enhance personal data privacy. Due to the scope and scale of the many detailed and granular proposals, this Part neither parses nor endorses them. Rather, this Part assesses how some of the most prominent approaches might impact low-income people and suggests possible improvements to the current privacy landscape. The political system is less responsive to low-income than wealthy Americans[324] and low-income Americans have less access to the

Fourth Amendment Privacy to Protect Associational Anonymity, 2014 U. ILL. J.L. TECH. & POL'Y 1 (2014); Tabatha Abu El-Haj, *Friends, Associates, and Associations: Theoretically and Empirically Grounding the Freedom of Association*, 56 ARIZ. L. REV. 53 (2014).

320. *See* Solove & Hartzog, *The New Common Law*, *supra* note 142, at 587.

321. Although this Article focuses on federal law due to its national scope, it is important to recognize that state laws also govern online privacy and big data collection. Partly due to congressional intransience, some state legislators and attorneys general have been particularly energetic in protecting and enforcing consumer privacy interests. *See* Danielle Keats Citron, *The Privacy Policymaking of State Attorneys General*, 92 NOTRE DAME L. REV. 747 (2016); Ganka Hadjipetrova & Hannah G. Poteat, *States are Coming to the Fore of Privacy in the Digital Era*, 6 LANDSLIDE 1, 12 (July/Aug. 2014); Paul M. Schwartz, *Preemption and Privacy*, 118 YALE L.J. 902, 917–18 (2009). For a list of relevant laws, *see State Laws Related to Internet Privacy*, NAT'L CONF. STATE LEGIS. (Jan. 5, 2016), https://perma.cc/FEE3-88VQ.

322. *See* BJ Ard, *The Limits of Industry-Specific Privacy Law*, 51 IDAHO L. REV. 607, 607 (2015) ("[T]he distinct features of online commerce . . . challenge discrete industry-specific laws.").

323. *See* Daniel J. Solove, *Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1880, 1880–81 (2013).

324. *See* MARTIN GILENS, AFFLUENCE AND INFLUENCE: ECONOMIC INEQUALITY AND POLITICAL POWER IN AMERICA 70–123 (2012). Gilens' research finds "a fairly strong association between policy outcomes and the preferences of the affluent, and weaker associations for the preferences of the middle class and the poor." *Id*. at 5.

justice system.[325] As a result, their voices are often silenced, despite their heightened vulnerability to privacy intrusions. Thus, our goal is to emphasize the privacy concerns facing low-income Americans in current policy discussions.

A. Notice and Choice

In the absence of broad statutory protections for big data, the FTC and the Obama Administration pushed self-regulation by business, largely through notice and choice offered to consumers.[326] This has been and remains the current governing privacy paradigm in the United States.[327] The goal of notice and choice is to provide consumers with information about a website's privacy policy, including its collection, use, and sharing practices, in order to allow the consumer to decide whether or not to use a certain website.[328] This model promotes the important value of self-autonomy, but assumes a fair contractual bargain between the website and the consumer, which is a myth for several reasons. Moreover, the assumptions underlying notice-and-choice are likely even more attenuated for under-resourced people.

To begin with, studies show that consumers do not understand website privacy notices because they are purposefully complex, lengthy, and jargon filled.[329] This can be particularly pernicious for less educated

325. *See generally* LEGAL SERVS. CORP., THE JUSTICE GAP: MEASURING THE UNMET CIVIL LEGAL NEEDS OF LOW-INCOME AMERICANS 6 (2017), http://www.lsc.gov/sites/default/files/images/TheJusticeGap-FullReport.pdf.

326. *See* Joel R. Reidenberg et al., *Privacy Harms and the Effectiveness of the Notice and Choice Framework*, 11 I/S: J.L. & POL'Y. FOR INFO. SOC'Y, 485, 486–87 (2015); Hoofnagle & Urban, *supra* note 10, at 1; Solove & Harztog, *The New Common Law, supra* note 142, at 592. For a recent description and endorsement of notice and choice, *see* FED. TRADE COMM'N, FTC ISSUES FINAL COMMISSION REPORT ON PROTECTING CONSUMER PRIVACY: AGENCY CALLS ON COMPANIES TO ADOPT BEST PRIVACY PRACTICES (2012), https://perma.cc/7PM6-ZV6K.

327. *See* Solove, *supra* note 323, at 1880–81 (describing structural and cognitive barriers to privacy self-management regimes). There are some indications that the FTC under the Trump Administration might narrow consumer protections by recognizing only privacy violations that result in harm, rather than a broader approach that gives consumers the right to opt out of sharing personal data with businesses regardless of whether or not harm would result from disclosure. *See* James R. Hood, *FTC's New Head Eyes "Harms-Based Approach" to Privacy Protection*, CONSUMER AFF. (Feb. 8, 2017), https://perma.cc/BJE9-MNRR.

328. *See* Robert H. Sloan & Richard Warner, *Beyond Notice and Choice: Privacy, Norms, and Consent*, 24 J. HIGH TECH. L. 370, 374 (2014).

329. *See* Reidenberg et al., *supra* note 326, at 491; Robert Sprague & Corey Ciocchetti, *Preserving Identities: Protecting Personal Identifying Information Through Enhanced Privacy Policies and Laws*, 19 ALB. L.J. SCI. & TECH. 91, 122–23 (2009); Sloan & Warner, *supra* note 328, at 391; Jan Whittington & Chris Jay Hoofnagle, *Unpacking Privacy's Price*, 90 N.C. L. REV. 1327,

consumers, who are also more likely to be low-income.[330] Indeed, recent surveys have supported earlier research indicating that adults with lower levels of education are more likely to falsely assume that when a company posts a privacy policy, it ensures the company will keep all of the information it collects on users confidential.[331] Even the most diligent consumer would lack the time to read the hundreds of privacy policies he or she might encounter in a single day.[332] And that time might be wasted given that companies often reserve the right to change their policies in the future without additional notice and consent, thus upsetting any initial bargain struck by a consumer.[333] Complicated privacy policies and complex privacy settings may be by design—if websites and social media applications require individuals to share information to make money, it may be in their best interest to keep individuals in the dark about their use of personal data. Moreover, even when a privacy policy is clear and comprehensible, consumers have no way to control how their data might be used downstream by third parties or later aggregated into a larger portfolio of information about them.[334] Moreover, a person who opts out of a certain website may not be aware that another person's disclosure

1357–58 (2012); Mark A. Graber, Donna M. D'Alessandro & Jill Johnson-West, *Reading Level of Privacy Policies on Internet Health Web Sites*, 51 J. FAM. PRAC. 642, 642 (2002).

330. Notice and choice is no barrier to discrimination, "[i]nstead, giving individuals notice and choice may simply perpetuate the growing gap between consumer 'haves' and 'have-nots' because the least sophisticated consumers remain least likely to protect themselves." Schmitz, *supra* note 139, at 1462–63. *See* Koenig & Rustad, *supra* note 209, at 616, 620, 627 (describing the lower levels of reading comprehension of low-income persons and how this renders privacy notices useless).

331. *See* Aaron Smith, *What Internet Users Know About Technology and the Web*, PEW RES. CTR. (Nov. 25, 2014), https://perma.cc/8M7R-ETSX.

332. *See* Reidenberg, *supra* note 326, at 492. Aleecia McDonald and Lorrie Cranor calculated that if all Americans were to read their privacy policies, the opportunity cost would be $781 billion. Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, I/S: J.L. & POL'Y FOR INFO. SOC'Y 544, 561 (2008).

333. *See* Sprague & Ciochetti, *supra* note 329, at 126; danah boyd & Eszter Hargittai, *Facebook Privacy Settings: Who Cares?*, 15 FIRST MONDAY 1, 2 (2010), https://perma.cc/SQG3-W7SY; Paul Ohm, *Branding Privacy*, 97 MINN. L. REV. 907, 909 (2013) (describing how such privacy lurches (or changes in promised policies) "disrupt long-settled expectations" and proposing that privacy policies become part of a company's brand that can only be changed with a new name).

334. *See* Reidenberg et al., *supra* note 326, at 492; Schmitz, *supra* note 139, at 1425; Natalie Kim, *Three's a Crowd: Towards Contextual Integrity in Third-Party Data Sharing*, 28 HARV. J. L. & TECH. 325, 327 (2014) ("privacy policies essentially remain a blunt instrument, giving users a binary option between sharing with none or sharing with all . . . ."); Clark D. Asay, *Consumer Information Privacy and the Problem(s) of Third-Party Disclosures*, 11 NW. J. TECH. & INTELL. PROP. 321, 324 (2013).

implicates them in ways of interest to data miners.[335]

Furthermore, notice and choice regimes are no guarantee of privacy. Numerous companies have failed to adhere to their stated privacy policies, failed to keep data secure, and retained data longer than what users deem as reasonable.[336] For all these reasons, self-regulation appears inadequate alone to protect consumer privacy interests. A major report in 2011 concluded that companies create self-regulation regimes and rules in secret, lack consumer representatives, only involve a fraction of covered industries, survive for only short periods, are underfunded, and lack the ability to enforce rules or maintain members.[337] Often, companies are unwilling to self-regulate along even the most basic privacy-protective principles.[338] Accordingly, scholars have argued that self-regulation can at times function as a charade providing cover to industry in lieu of substantive regulation.[339]

In addition, notice and choice simply cannot protect against certain harms, such as failure to adhere to the terms of a notice, negligent security practices, or wrongful retention of personal data.[340] In other words, notice and choice does not protect against broken promises. Other legal protections are required to enforce any bargain struck by consumers.

Even where legal protections exist, such as with FCRA, we need to be mindful of the limits of private litigation when it comes to low-income individuals. Most low-income consumers will never know that they have been subject to adverse action based on personally identifiable information. Even if they become aware of a privacy violation, they lack the resources to hire a lawyer, and civil legal services satisfy only 14% of the legal needs of low-income litigants.[341] Statutory damages for privacy violations under FCRA are low.[342] Many agreements that bind consumers

335. Reidenberg et al., *supra* note 326, at 495

336. *Id*. at 521–23.

337. ROBERT GELLMAN & PAM DIXON, WORLD PRIVACY FORUM, MANY FAILURES: A BRIEF HISTORY OF SELF-REGULATION IN THE UNITED STATES 6 (2011), https://perma.cc/WE6R-SDGV.

338. *See* Natasha Singer, *Consumer Groups Back Out of Federal Talks on Face Recognition*, N.Y. TIMES BITS BLOG (June 16, 2015, 12:10 AM), https://perma.cc/VR6C-GQD2.

339. *See* Chris Hoofnagle, *Can Privacy Self-Regulation Work for Consumers?*, TAP BLOG (Jan. 26, 2011), https://perma.cc/2KS7-7NPL.

340. Reidenberg, et al., *supra* note 326, at 521. In addition, notice and choice is nearly impossible with regard to the Internet of things, which is increasingly capturing personally identifiable information. *See* Meg Leta Jones, *Privacy Without Screens & The Internet of Other People's Things*, 51 IDAHO L. REV. 639, 640 (2015).

341. LEGAL SERVICES CORPORATION, *supra* note 325, at 6.

342. *See* Austin H. Krist, Note, *Large-Scale Enforcement of the Fair Credit Reporting Act and the Role of State Attorneys General*, 115 COLUM. L. REV. 2311, 2321–22 (2015). Under FCRA, potential

push them into arbitration, rather than court, and arbitration is typically a process too expensive for an individual to pursue. In short, there is almost no incentive or ability for a private consumer to bring a data privacy claim. For all these reasons, notice and choice—even an improved version of notice and choice—is not sufficient on its own to protect consumer privacy, and it is particularly problematic for low-income Americans.

B. Digital Literacy

Given the well-known shortcomings in notice and choice regimes, some advocates have argued for greater digital literacy so that people can better avail themselves of notice and choice protections. *Digital literacy* refers to the vast array of "technical, cognitive, and sociological skills" that individuals need "in order to perform tasks and solve problems in digital environments."[343] The term suffers from a lack of precision, and is used to refer to: assessing information credibility, mastery of certain technical skills, knowledge of computer hardware and peripherals, familiarity with software interfaces, social-emotional awareness of digital environments, and so forth.[344] Digital literacy programs in which individuals are taught some of these skills, are often championed as potential solutions for privacy violations, particularly when those violations could have been prevented by self-help behaviors by individuals. Some programs explicitly teach privacy practices and behaviors, such as how to change privacy settings on social media, while others generally discourage people from sharing online. Some researchers postulate that digital literacy programs that improve people's Internet skills may "support, encourage, and empower users to undertake informed control of their digital identities."[345]

One possible interpretation of the results presented in Part II of this paper is that privacy education, or increased privacy literacy, may help

compensatory damages for an individual plaintiff are low. A violation results in a maximum award of $1000, and then only for willful noncompliance, which is hard to prove and which courts rarely find. 15 U.S.C. §1681n (2012).

343. Yoram Eshet-Alkalai, *Digital Literacy: A Conceptual Framework for Survival Skills in the Digital Era*, 13 J. EDUC. MULTIMEDIA & HYPERMEDIA 93 (2004).

344. Yoram Eshet Alkali & Yair Amichai-Hamburger, *Experiments in Digital Literacy*, 7 CYBERPSYCHOLOGY & BEHAV. 421 (2004); DIGITAL LITERACIES: CONCEPTS, POLICIES AND PRACTICES (Colin Lankshear, Michele Knobel & Michael Peters eds., 2008).

345. Yong Jin Park, *Digital Literacy and Privacy Behavior Online*, 40 COMM. RES. 215, 217 (2013).

solve the problems of low-income Internet users. Our empirical data suggests that when compared to higher-earning people, low-income users are less confident in their ability to manage their privacy settings, understand privacy policies, and find tools and strategies that would help them protect their personal information online. While privacy literacy may help people to feel more effective at managing their information, and may reduce some of their exposure to data collection, it is not a sufficient solution to the problems that the poor face with regards to privacy and information flow.

First, as noted above,[346] most users find privacy settings and privacy policies abstruse and difficult to understand. Second, harms related to data collection and dissemination are institutional problems that cannot be fully ameliorated through individual use of extant privacy tools, even if data-related vulnerabilities of the poor are exacerbated by lack of privacy-protective behaviors (e.g., adjusting privacy settings). Even if individuals chose to abstain from social media entirely, their personal information would still be collected from a myriad of sources including, but not limited to government, public, and court records; motor vehicle and driving records; recorded mortgages and tax assessments; catalog and magazine subscriptions; store loyalty cards; warranties; the US Census; voter registration information; financial records; and others.[347] Simply moving through a city, using public transportation, or driving on toll roads creates a data trail that is almost impossible to avoid.[348] Putting the burden on individuals alone to protect their privacy ignores the multitude of ways in which information is collected, tracked, and aggregated in ways that individuals cannot control, and of which they may not be aware.

Most importantly, emphasizing privacy literacy in the absence of other possible reforms shifts the responsibility for privacy protection to the individual. This suggests that if an individual's privacy is violated, it is because she did not protect it adequately. This places the fault on the individual rather than the person or organization that violated her privacy. This rhetoric of individual responsibility is also found in the public debate over social services, in which poverty is seen as resulting from a series of bad decisions made by individuals, rather than a systemic problem.[349]

346. *See supra* notes 329–335 and accompanying text.

347. FED. TRADE COMM'N, *supra* note 57, at 11–15.

348. Kessler, *supra* note 69.

349. *See* Ronen Shamir, *The Age of Responsibilization: On Market-Embedded Morality*, 37 ECON. & SOC'Y 1 (2008); Martha Poon & Helaine Olen, *Does Literacy Improve Finance?*, 24 PUB. UNDERSTANDING SCI. 272 (2015).

This, in turn, justifies some of the wide-reaching privacy violations detailed earlier in this document that contribute to the dehumanization of the poor, such as welfare home visits, mandatory drug testing, and the like in order to receive social services.[350] While privacy literacy programs provide an important foundation for improving consumer awareness of privacy-enhancing tools and strategies, they do little to restrain the larger ecosystem of data brokers, consumer profiling and government surveillance discussed in this paper.

C. Due Process

Several scholars have argued that due process norms should apply to big data applications, especially those that impact people's access to credit, employment, education and other life needs.[351] Procedural due process is enshrined in the Constitution and gives individuals the right to notice and a hearing regarding governmental actions that threaten to take away their life, liberty, or property.[352] While due process rights only attach to government actions, the quasi-governmental nature of how private entities are determining public access to valuable opportunities makes the values of due process—including "transparency, accuracy, accountability, participation, and fairness"—appealing in a big data context.[353] To enforce these values, scholars have recommended that algorithmic formulas be made public and that people who are adversely impacted have the right to challenge inaccurate information about them or unfair outcomes.

As in the governmental context, "data due process" may be a limited remedy for low-income individuals. To begin with, most people are not aware of the extent to which their data are being gathered and aggregated to make decisions about them.[354] Recognizing this, advocates of this approach suggest that every adverse decision against an individual be

350. *See* Gilman, *Class Differential*, *supra* note 1, at 1391–92; Nancy Fraser & Linda Gordon, *A Genealogy of Dependency: Tracing a Keyword of the U.S. Welfare State*, 19 SIGNS 309 (1994); Jerry Watts & Nan Marie Astone, *Review: The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 by 104th Congress of the United States*, 26 CONTEMP. SOC. 409 (1997).

351. *See* Crawford & Schultz, *supra* note 48; Citron & Pasquale, *supra* note 152.

352. Goldberg v. Kelly, 397 U.S. 254 (1970).

353. Citron & Pasquale, *supra* note 152, at 19–20.

354. *See* Neil M. Richards & Jonathan H. King, *Big Data Ethics*, 49 WAKE FOREST L. REV. 393, 393, 414 (2014). "Institutions like data brokers, often without our knowledge or consent, are collecting massive amounts of data about us they can use and share in secondary ways that we do not want or expect." *Id*. at 421.

revealed, along with an audit trail that reveals the basis of automated decisions.[355] When it comes to governmental benefits, a notice and explanation of denial is fairly straightforward because these determinations are usually made in the context of objective eligibility criteria. However, such requirements would engender a radical reshaping of the workplace, higher education, and related settings if every disappointed applicant were entitled to a reason for their rejection along with a dossier of algorithmic formulas that might have contributed to the rejection, especially given that big data often interacts with individual discretion in these situations.

Moreover, due process has proved a mixed revolution when it comes to public benefits.[356] While it has provided low-income people with a forum to assert their rights, its shortcomings include a lack of lawyers to enforce those rights, an adversarial rather than problem-solving approach, an often demeaning and confusing process, and a masking of systemic injustice through the framework of individual fair hearings. Thus, in the big data context, it would be essential for any due process regime to ensure that low-income people have a voice in designing systems for transparency and accountability, that their interests are represented by enforcement entities, and that enforcement involves systemic review of algorithmic processes rather than reliance on individual complaints.

D. Comprehensive Consumer Privacy Legislation

The United States' sectoral approach to privacy is often compared to the European Union's (EU) comprehensive, *ex ante* approach to privacy.[357] In the EU, data collection must be limited in scope and retention, and data is subject to consumer consent, review, and correction.[358] Data subjects also have the right not to be subject to

355. Crawford & Schultz, *supra* note 48, at 125.

356. *See* Jason Parkin, *Adaptable Due Process*, 160 U. PA. L. REV. 1309, 1331 (2012) (summarizing critiques of fair hearing rights, including that "[f]air hearings do not address barriers that prevent eligible individuals from applying for welfare in the first instance"); Rebecca E. Zietlow, *Giving Substance to Process: Countering the Due Process Counterrevolution*, 75 DENV. U. L. REV. 9, 26 (1997) ("formal procedural rights may hurt rather than help poor people because they serve to mask substantive injustice"); Michael Herz, *Parallel Universes: NEPA Lessons for the New Property*, 93 COLUM. L. REV. 1668, 1710 (1993) (due process in public benefits regimes has "contributed to routinization, alienation, and abuse").

357. *See generally* James Q. Whitman, *The Two Western Cultures of Privacy: Dignity Versus Liberty*, 113 YALE L.J. 1151 (2004).

358. For a summary of EU data protections, *see* Marc Rotenberg & David Jacobs, *Updating the Law of Information Privacy: The New Framework of the European Union*, 36 HARV. J. L. & PUB.

decisions with legal effects generated from automated processing.[359] In 2016, the EU passed the General Data Protection Regulation, which provides even stronger privacy protections, including a "right to be forgotten," strict consent requirements that place the burden of proof on the collector, and increased sanctions for violations.[360] The Regulation also tightens restrictions on the processing of particularly sensitive information, such as race, political opinions, and religion.[361] In addition, the Regulation gives people the right not to be subject to decisions solely based on algorithms, including profiling, and to contest such decisions.[362] Almost all commentators acknowledge that the United States, with its emphasis on personal liberty and corporate innovation, is very unlikely to adopt an omnibus privacy statute in the style of the EU, with its emphasis on personal dignity.[363]

Accordingly, a range of policymakers, business leaders, and privacy advocates have argued for comprehensive privacy legislation in the more limited sphere of online consumer activity.[364] For instance, in 2012, the Obama White House issued a proposed Consumer Privacy Bill of Rights based on seven principles of individual control: transparency, respect for context, security, access, accuracy, focused collection, and

---

POL'Y 605, 615–622 (2013).

359. *Id*. at 619.

360. *See* Stephen Gardner, *EU Parliament Finalizes Landmark Data Privacy Reg*, BLOOMBERG LAW (April 28, 2016), https://perma.cc/L6XT-K873. For detailed descriptions of the new regulation, *see* Rotenberg & Jacobs, *supra* note 358, at 630–36; Paul M. Schwartz, *The EU-U.S. Privacy Collision: A Turn to Institutions and Procedures*, 126 HARV. L. REV. 1966, 1992–2003 (2013).

361. *See* Schwartz, *Privacy Collision, supra* note 360, at 1996. For a comparison of U.S. and EU perspectives with regard to protections against discrimination, *see generally* Raymond Shih Ray Ku, *Data Privacy as a Civil Right: The EU Gets It?*, 103 KY. L.J. 391 (2014–2015).

362. See Bryce Goodman & Seth Flaxman, *European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation,"* OXFORD INTERNET INST., https://arxiv.org/pdf/1606.08813.pdf (describing Article 22 and the challenges and opportunities inherent in this new provision.). The EU approach is not a panacea, as it is complex and raises challenging enforcement issues. *See* Michiel Rhoen, *Beyond Consent: Improving Data Protection Through Consumer Protection Law*, 5 INTERNET POL'Y REV. (2016), https://policyreview.info/articles/analysis/beyond-consent-improving-data-protection-through-consumer-protection-law.

363. There is a debate over whether the federal government or the states are better sites for privacy law. *Compare* Schwartz, *Preemption and Privacy*, *supra* note 330, at 916-17 (arguing against comprehensive federal regulation), *with* Patricia L. Bellia, *Federalization in Information Privacy Law*, 118 YALE L.J. 868 (2009) (arguing that Schwartz overstates the risks of omnibus federal regulation). Several states are more protective of online personal data.

364. *See* Schwartz, *Preemption and Privacy*, *supra* note 321, at 904 (describing calls for comprehensive privacy legislation and arguing such an approach is misguided).

accountability.[365] This set of baseline protections would give consumers the right to determine what data companies collect about them and how those data are used, although it envisions a combination of industry self-regulation and government enforcement.[366] Relatedly, the FTC recommended in 2014 that Congress enact legislation to regulate the data broker industry so that consumers would know about the industry's activities and have access to information held about them by data brokers.[367] In recent years, Congress has considered, but failed to pass, a variety of data privacy bills.

Almost any of these proposals would enhance data privacy while providing companies with greater guidance as to their obligations. For low-income consumers in particular, the effectiveness of these laws would hinge upon the existence of meaningful sanctions, rigorous oversight by governmental or third-party entities of data collection and processing, clear notice and consent policies optimized for mobile devices, and a legal commitment to identifying and ameliorating harmful and unjustified disparate impacts.

E. Areas for Further Research

In light of these substantial challenges, there continues to be a great need for interdisciplinary research to deepen our understanding of the class differential in privacy vulnerability. This Article has sought to provide a foundation for further inquiry, describing a matrix of overlapping vulnerabilities that low-income communities face in the big data era. As this analysis has illustrated, there are both considerable gaps in current legal protections and the information available to the public to assess the fairness of the rapidly evolving methods of automated decision-making in employment, education, and law enforcement contexts. These gaps are especially acute when considering the role of social media and network-based assessments, the technical details of which are largely opaque and ripe for further investigation by researchers.

Additional studies could further explore the connection between a

365. The White House, *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy*, 4 J. PRIVACY & CONFIDENTIALITY 95 (2012), https://perma.cc/8A4U-Y9SP.

366. *Id*. at 104. Both companies and privacy advocates expressed complaints about the Consumer Bill of Rights for going too far and not far enough respectively. *See* Brendan Sasso, *Obama's Privacy Bill of Rights Gets Bashed from All Sides*, ATLANTIC (Feb. 27, 2015), https://perma.cc/P9LG-R9GV.

367. FED. TRADE COMM'N, *supra* note 57, at 49.

reliance on mobile access among low-income Internet users and increased vulnerability to data broker profiling. Looking specifically at social media exposure, deeper analyses of what kinds of content are made public by default across different mobile versions of social media platforms would build upon prior work examining changes in default sharing settings over time across popular social network sites.[368] Future empirical studies could also explore the public's awareness of social media monitoring in these environments and how this varies across socioeconomic groups.

Looking ahead, it is likely that data-driven innovation and its associated economic efficiencies will continue to outpace the implementation of legal constraints to prevent potentially biased or unfair decision making practices for low-income communities. Ultimately, the complexity of the current technological and legal environments will increasingly require researchers and advocates to employ a range of methods to ensure that the interests of those who are most vulnerable to privacy-related harms are not overlooked or written off as necessary collateral for realizing the benefits of the big data era.

CONCLUSION

We live in an age of increasing income and wealth inequality with an economic gulf between the rich and poor and a stagnating middle-class. Data-driven systems could play a role in reversing these trends by expanding access to education, employment, and justice for marginalized populations. However, they might conversely contribute to a widening of that gap by providing a ready avenue to prey on the vulnerabilities of low-income people, or to exclude them from opportunities due to biases entrenched in algorithmic decision-making tools. Historically, poor people have faced much greater surveillance than their wealthier counterparts, and anti-poverty advocates are rightfully concerned that the digital world will replicate, if not reinforce, both covert and overt patterns of surveillance. For their part, low-income Americans express greater concerns regarding data collection in a variety of contexts, but they are more likely to access the Internet from less secure mobile devices, and to report lower usage of privacy settings and protective strategies. The three case studies in this Article highlight some forms of potentially harmful discrimination that low-income Americans might face when data analytics are used in hiring,

368. *See supra* note 96.

college admissions, and law enforcement. In each setting, low-income Americans face not only adverse inferences drawn based on their personally identifiable information (which often is erroneous), but also those drawn from their social media and demographic networks. Existing law provides little recourse, as it mostly pre-dates the proliferation of the Internet and favors business rather than consumer interests through a self-regulatory regime. Given that the political system tends to be less responsive to the needs of low-income Americans and that they often lack access to the justice system, it is imperative that policy discussions around digital privacy increasingly include the voices and perspectives of low-income people.

APPENDIX: SUMMARY OF SURVEY METHODS

The survey on Privacy and Security Experiences of Low-Socioeconomic Status Populations, sponsored by the Data & Society Research Institute, obtained telephone interviews with a nationally representative sample of 3000 adults ages eighteen and older living in the United States. Interviews were completed in both English and Spanish, according to the preference of the respondent. The survey was conducted by Princeton Survey Research Associates International (PSRAI). The interviews were administered by Princeton Data Source from November 18 to December 23, 2015. A combination of landline and cell phone random-digit dial (RDD) samples was used to reach respondents regardless of the types of telephone they have access to. Both samples were disproportionately stratified to target low-income households. A total of 1050 interviews were conducted with respondents on landline telephones and 1950 interviews were conducted with respondents on cellular phones, including 1193 who live in a household with no landline telephone access.

Statistical results are weighted to correct for the disproportionate sample design, the overlapping landline and cell sample frames and disproportionate non-response across demographic groups that might bias results. The final weighted total sample is representative of all adults ages eighteen and older living in the United States. The margin of sampling error for the complete set of weighted data is ±2.7 percentage points.

One inherent limitation of survey research is what is known as "social desirability bias." The Pew Research Center's methodology experts describe this as people's "natural tendency to want to be accepted and

liked," which may in turn lead people to report inaccurate answers to questions that probe on sensitive subjects.[369] In addition, it may be the case that surveys underreport privacy concerns and sensitivities because those who respond to surveys are generally more comfortable sharing information about themselves. In terms of mode effects, however, recent studies have suggested that there are not meaningful differences between telephone and online surveys with respect to the degree people express worries about computers and technology "being used to invade privacy."[370]

---

369. *See Questionnaire Design,* PEW RES. CTR. (Apr. 2017), https://perma.cc/3VNR-MZFF .

370. *See From Telephone to the Web: The Challenge of Mode of Interview Effects in Public Opinion Polls*, PEW RES. CTR. (MAY 13, 2015), https://perma.cc/XJB9-WU6E.