Ohio Northern University Law Review

Volume 42 | Issue 2 Article 5

Consumer Protection in an Era of Big Data Analytics

David C. Vladeck

Consumer Protection in an Era of Big Data Analytics

DAVID C. VLADECK*

For most of human history, the idea that machines would evolve to the point that they would assist humans in decision-making was the stuff of science fiction. Science fiction writers have long been of two minds about what might happen to humans if machines could actually "think." One vision was decidedly Utopian. Machines helped humans solve problems beyond their ability to resolve, and partnered with humans to carry out the run-of-the-mill, mundane work humans perform, freeing us to pursue higher callings.

To a large extent, that vision is being realized. Driverless cars are tested on our roads today. Autopilot devices fly the planes we travel on across continents. Surgical robots perform operations too delicate for human hands. High-speed computers make stock and commodity trading decisions in microseconds, by-passing the delay of human intervention. And, as forecasted, industrial robots now perform much of the routinized, assembly-line work that marked the Industrial Revolution.

The growing dominance of robots in the manufacturing sector, which has caused enormous economic dislocation and triggered the erosion of our middle class, also feeds into the darker vision of the potential competition between machines and humans. Perhaps the best science fiction expresses the dystopian view that, at some point, autonomously thinking machines will turn on humans and pose an existential threat to the human race. Indeed, more than a few artificial intelligence experts, including Stephen Hawkins and Elon Musk, worry that that point is not long off,[1] even though no one would contend that today's machines have reached a level of beyond-human-intelligence, or "singularity," that poses that sort of risk.

---

* Professor of Law, Georgetown University Law Center, and faculty director of Georgetown's Center on Privacy and Technology. This Essay draws on the author's experience as Director of the Federal Trade Commission's Bureau of Consumer Protection from 2009 through 2012. An earlier version was presented at the 39th Annual Ohio Northern University Law Review Symposium as the Carhart Lecture to students and faculty at ONU Pettit College of Law. Special thanks to Joshua Lanphear, Editor-in-Chief of the Law Review, for his top-notch editorial assistance.

1. *See* Future of Life Institute, *An Open Letter: Research Priorities for Robust and Beneficial Artificial Intelligence*, http://futureoflife.org/ai-open-letter/ (last visited Mar. 14, 2016); Future of Life Institute, *Autonomous Weapons: An Open Letter From AI & Robotics Researchers* (July 28, 2015), http://futureoflife.org/ai-open-letter/. *See generally* David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence*, 89 WASH. L. REV. 117 (2014) (discussing the impact of the introduction of autonomous machines into the marketplace).

This essay focuses on non-existential but nonetheless clear and present threats posed by the ubiquitous collection of personal data and the use of big data analytics to make highly consequential decisions about each of us. These decisions are based on opaque and ever-evolving machine learning algorithms that process deep reservoirs of data, some of which may be incomplete or inaccurate, with little oversight by humans and little, if any, opportunity to challenge those decisions.[2]

In this new world, individuals are no longer judged on what they have done. Instead, we are judged based on inferences or correlations drawn by algorithms and other analytic techniques that determine whether we should be categorized in certain ways that may make us appear to be poor credit or insurance risks, unsuitable candidates for employment or admission to schools, or otherwise unacceptable or less desirable to companies and institutions that provide important goods and services.[3] Decisions that matter will now be based on correlations, not hard facts.[4] For companies that deal with thousands of consumers, using correlations in this reductive way to make decisions may make sense.[5] The law of large numbers means that, if the algorithm is reasonably accurate, then in general the companies will be making the right call.[6] To the extent the algorithm misfires, the unfairly characterized consumer is simply collateral damage.[7]

But to the consumer who has been wrongly categorized, that mis-categorization may feel like arbitrariness-by-algorithm. Let me be clear: The fact that decision-by-algorithm may be less than perfect is not to condemn the enterprise. Far from it. Using big data analytics to improve decision-making is, in some respects, an important step forward.[8] After all, human decision-making is far from error-free. People often make imperfect decisions for a variety of reasons, including incomplete or inaccurate

2. FED. TRADE COMM'N, BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION? UNDERSTANDING THE ISSUES 5-12 (Jan. 2016), https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf [hereinafter BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?].

3. Solon Barocas & Andrew D Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. at 7-8 (forthcoming 2016).

4. *Id.*

5. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 5-7.

6. PAM DIXON & ROBERT GELLMAN, WORLD PRIVACY FORUM, THE SCORING OF AMERICA: HOW SECRET CONSUMER SCORES THREATEN YOUR PRIVACY AND YOUR FUTURE 41 (Apr. 2, 2014) http://www.worldprivacyforum.org/wp-content/uploads/2014/04/WPF_Scoring_of_America_April2014_fs.pdf. The law of large numbers is a principle of probability and statistics that states that as a sample size grows, its mean will get closer and closer to the average of the whole population. *See* Richard Routledge, *Law of Large Numbers: Statistics*, ENCYCLOPAEDIA BRITANNICA, http://www.britannica.com/science/law-of-large-numbers (last visited Mar. 14, 2016).

7. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 26-27.

8. *Id.* at 1-2.

information, poor decisional tools, or irrational bias. But the decision-by-algorithm process is no panacea for flawed human decision-making.[9] Algorithms may also be imperfect decisional tools. Algorithms themselves are designed by humans, leaving open the possibility that unrecognized human bias may taint the process.[10] And algorithms are no better than the data they process, and we know that the much of that data may be unreliable, outdated, or reflect bias.[11] Machines cannot completely eliminate human error.

For these reasons, algorithmic decision-making demands transparency, meaningful oversight, and procedures to remediate decisions that adversely affect individuals who have been wrongly categorized by correlation.[12] At the very least, companies and other institutions that employ algorithmic decision-making must take steps to ensure that their algorithms do not end up making determinations about individuals based on categories that society has decided—by law or ethical norms—not to use, such as race, ethnic background, gender, sexual orientation, and religious beliefs.[13] This essay explores some of the risks posed by the use of big data analytics for decision-making and proposes some ways to mitigate those risks.

I. BIG DATA DECISION-MAKING

Algorithmic decision-making is the process of using of mathematical formulas (algorithms) to sort through data (inputs) in order to answer certain questions.[14] Some questions are straightforward (*e.g.*, the conversion of inches to meters), and these well-defined problems are generally solved with straightforward formulas.[15] But algorithms can also be used to solve complex, ill-defined problems that have to take multiple variables and massive amounts of data into account and do not necessarily yield a single, "right" answer.[16] Complex algorithms are now in wide use, many of which are familiar to us, such as sorting through huge data sets to improve weather forecasting, safeguarding computer networks from intrusion, and predicting which middle school students may struggle in high school so educational

9. Barocas & Selbst, *supra* note 3, at 3.
10. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 28.
11. *Id.*
12. *Id.* at 27-30.
13. *See* Barocas & Selbst, *supra* note 3, at 3-4.
14. FAIR ISAAC, FAIR ISAAC CORPORATION 2005 ANNUAL REPORT: DECISIONS MADE SIMPLE 20 (2005), http://library.corporate-ir.net/library/67/675/67528/items/177564/FIC_AR2005.pdf.
15. DIXON & GELLMAN, *supra* note 6, at 28 n.33.
16. *Id.* at 38-39.

interventions can take place earlier and thus be more effective.[17] All of that, and more, adds measurable value to society.[18]

But algorithms are also used for purely commercial purposes, and the social value of these uses is, at least in some instances, contestable.[19] For example, marketers use algorithms to sort through enormous amounts of personally-identifiable data, both to segment individuals based on certain characteristics and interests, and to make data-driven judgments about which individuals might be interested in buying particular products, such as cars or laundry detergent.[20] When viewed in that light, big data analytics may seem benign.

Other uses, however, may be more problematic. Algorithms may also be used to make marketing judgments that might be highly consequential.[21] For instance, algorithms are increasingly being used to determine whether an individual is credit-worthy and if so, to what extent; whether owning a motorcycle makes a person a poor credit, insurance, or employment risk; or whether patronizing certain stores suggests that a person may be tardy in paying a car loan.

Before the era of "big data," the algorithms used to make these determinations were relatively straightforward and relied on discrete data sets.[22] Consider one example. The most common credit scoring methodology, developed by the Fair Isaac Company ("FICO") decades ago and still in use today, considers just five variables in determining an individual's credit score: the individual's (1) payment history, (2) outstanding debt, (3) length of credit history, (4) pursuit of new credit, and (5) debit-to-credit ratio.[23] To be sure, FICO continues to fine-tune the values assigned to each of these factors, but at least historically the process by which it arrives at a credit score is relatively transparent.[24] Changing any

17. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 6-7 (detailing a wide range of uses for big data algorithms).

18. *Id.* at 5.

19. *Id.* at 9-11.

20. *Id.* at 5.

21. *Id.* at 9-11; *see* Ron Lieber, *American Express Kept a (Very) Watchful Eye on Charges*, N.Y. TIMES (Jan. 30, 2009), http://www.nytimes.com/2009/01/31/your-money/credit-and-debit-cards/31 money.html?pagewanted=all; Natasha Singer, *Mapping, and Sharing, the Consumer Genome*, N.Y. TIMES (Jun. 16, 2012), http://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html?pagewanted=all&_r=0.

22. *See* Barocas & Selbst, *supra* note 3, at 3; *see also* Mikella Hurley & Julius Adebayo, *Credit Scoring in the Era of Big Data*, __ YALE J. OF L & TECH. (forthcoming 2016) (manuscript on file with the author).

23. ROBINSON + YU, UPTURN, KNOWING THE SCORE: NEW DATA, UNDERWRITING, AND MARKETING IN THE CONSUMER CREDIT MARKETPLACE: A GUIDE FOR FINANCIAL INCLUSION STAKEHOLDERS 9 (Oct. 29, 2014), https://www.teamupturn.com/static/files/Knowing_the_Score_ Oct_2014_v1_1.pdf; Robert A. Avery et al., *An Overview of Consumer Data and Credit Reporting*, 89 FED. RESERVE BULL. 47, 47, 56-73 (Feb. 2003).

24. ROBINSON + YU, *supra* note 23, at 9.

one of the five variables would have a predictable effect on the score.[25] But FICO's historic approach is imperfect. To work, it needs reliable information on each of the five factors. As a result, FICO's first-generation credit scoring tool cannot be applied to new entrants in the credit market, who do not have the credit history that is crucial to FICO's traditional methodology.[26]

In the past, the marketplace punished those whose file was too "thin" to have a credit score by either denying them access to credit or offering them credit on less than favorable terms.[27] FICO understood the limitations of its approach, but it was constrained both by the limits on the information it could obtain from its financial partners and by its inability to store and process substantial quantities of data.[28] It may be hard for younger generations to understand this, but as recently as a decade ago, acquiring and storing data was costly, and big data analytics was the province of only the largest enterprises.

These constraints no longer exist. Over the past decade, the ability to collect, store, and analyze data about identifiable individuals has skyrocketed as quickly as the cost of doing so has plummeted.[29] Companies in the business of brokering information—now referred to as "data brokers"—have access to vast quantities of information about our online and offline activities, and assemble massive databases on virtually every person in the United States.[30] Data brokers sweep up all of the personally identifiable information made public by federal, state, and local governments, including information on property ownership, voter registration (*e.g.*, name, date of birth, address, and party affiliation), motor vehicle and driving records, and more.[31] They also collect personally-identifiable information from other public sources, such as press reports, social media sites and blogs, and other internet postings.[32] And data brokers buy information from other commercial sources—including financial

25. *Id.*

26. *Id.* at 10. FICO now has multiple scoring algorithms to take care of this "thin file" problem. *Id.* at 11-12.

27. *See id.* at 10-12.

28. *See* ROBINSON + YU, *supra* note 23, at 9-12.

29. DIXON & GELLMAN, *supra* note 6, at 15-17; *see* FED. TRADE COMM'N, REPORT TO CONGRESS UNDER SECTION 319 OF THE FAIR AND ACCURATE CREDIT TRANSACTIONS ACT OF 2003, (Dec. 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.

30. FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 8 (May 2014), https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf [hereinafter DATA BROKERS: A CALL FOR TRANSPARENCY]; *see* ROBINSON + YU, *supra* note 23, at 16.

31. DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 1-2; *see* ROBINSON + YU, *supra* note 23, at 16.

32. *Id.*

institutions, commercial entities, companies with loyalty card programs, customer lists from internet sites where consumers log in or register to obtain services, health information from consumers' fitness apps, information about consumers' browsing history from online advertising networks, and much more.[33] This data is then aggregated into a single, searchable database.[34]

Make no mistake, there is little question that the major data brokers know more about each of us than say, for example, the National Security Agency, the Internal Revenue Service, the Social Security Administration, or any other governmental institution.[35] Indeed, government agencies routinely buy information from data brokers to round out their own profiles.[36]

And we, of course, are partners in the corporate acquisition of our personal information. Take smartphones for example. Smartphones are wonderful tools, but they are also highly efficient information-sharing devices, passing along geolocation information constantly, not just to service providers, but to Google or Apple, app providers, and ultimately to companies that serve ads and conduct data analytics.[37] Wear a health tracking device? Those devices also collect highly personal information and generally share that information with a wide range of service providers and analytical companies.[38]

Acxiom, the largest data broker, acknowledges that it has an average of over 3,000 pieces of information on nearly every U.S. consumer.[39] Acxiom's data would likely include everything from an individual's salary

33. *Id.*

34. DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 1-2; *see* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1748 (2010) (referring to massive, aggregated data-bases as potential "database[s] of ruin" because of the likelihood that they contains highly sensitive data).

35. *See id.* at 11-14 (data brokers get information on consumers from more than just government entities).

36. *See e.g.*, *Industry: Government*, TRANSUNION, https://www.transunion.com/industry/government (last visited Mar. 14, 2016); *Government Need: Leverage Analytics*, EQUIFAX, http://www.equifax.com/government/leverage-analytics (last visited Mar. 14, 2016); *Public Sector Services*, EXPERIAN, http://www.experian.com/public-sector/public-sector-consulting.html (last visited Mar. 14, 2016); GENERAL SERVICES ADMINISTRATION FEDERAL SUPPLY SERVICE, AUTHORIZED SUPPLY SCHEDULE PRICE LIST (Sept. 1, 2011-Aug. 31, 2016), https://www.gsaadvantage.gov/ref_text/GS35F0604S/GS35F0604S_online.htm.

37. Comments of Alvaro Bedoya & David Vladeck, Center for Privacy & Technology at Georgetown Law, on "Big Data and Consumer Privacy in the Internet Economy," Docket No. 14514424-4424-01, to National Telecommunications & Information Administration (Aug. 5, 2014), http://www.law.georgetown.edu/academics/centers-institutes/privacy-technology/publications-filings/upload/8-5-14-Bedoya-and-Vladeck-Comment-FINAL.pdf.

38. Lisa Eadicicco, *A New Wave of Gadgets Can Collect Your Personal Information Like Never Before*, BUS. INSIDER (Oct. 9, 2014), http://www.businessinsider.com/privacy-fitness-trackers-smartwatches-2014-10.

39. DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 8.

information, the value of one's home, the make and model of one's car, driving history, educational experience, family relationships, family medical history, alcohol and drug use, religious affiliation, purchasing history, hobbies, political affiliation, and so on.[40] Datalogix, a data broker that follows consumer spending, provides businesses with marketing data on almost every U.S. household and on more than one trillion dollars in consumer transactions.[41] Equifax, another large data broker, collects salary data for well over one-third of our nation's workforce.[42] Tower Data Services (formerly Rapleaf), a specialized data broker, has an email address for most Americans, and those addresses are often linked to other personally-identifiable information.[43] These examples reflect just a tiny portion of the data broker industry.

The velocity of data collection will continue to increase as networked data collection extends into what were once our most private spaces—our homes and our cars. As a result of the Internet of Things, homes, cars, and even appliances will soon collect highly personal information about our lives and share that information with service providers, and almost certainly, data brokers.[44] Google's Nest will embed internet-enabled sensors in homes to manage thermostats, smoke and carbon dioxide detectors, and security cameras—all of which will keep families safe and comfortable.[45] Surely this a step forward. But all of this information will be collected by Google, so Google will know not only when you or your family members are at home, but also when you are sleeping, what rooms you tend to occupy, and when you are at work or on vacation.

The Internet of Things will also bring us internet-enabled household appliances.[46] For instance, the next refrigerator you buy will likely be able to track a household's use of core products and let the grocery store know when, Heaven forbid, you are nearly out of beer. Will it also order chips and salsa as a default, unless you say no? It may. And where else will that data go? Even conversations within the home are now fair game.

---

40. *See id.* at 11-14 (detailing the methods data brokers generally use to obtain information and the types of information they receive).

41. *Id.* at 8.

42. *See* Red Tape, *Exclusive: Your Employer May Share Your Salary, and Equifax Might Sell that Data*, NBC NEWS (Jan. 30, 2013) http://www.nbcnews.com/technology/exclusive-your-employer-may-share-your-salary-equifax-might-sell-1B8173066.

43. DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 9.

44. FED. TRADE COMM'N, STAFF REPORT, INTERNET OF THINGS: PRIVACY &SECURITY IN A CONNECTED WORLD 1-2 (January 2015 https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf [INTERNET OF THINGS].

45. Michael McCole, *How to Make Nest's Thermostat Your Smart-Home Hub*, WIRED (Feb. 10, 2016), http://www.wired.com/2016/02/iotcookbook-nest/.

46. INTERNET OF THINGS, *supra* note 44, at 1-2.

Televisions now have the capacity to listen to your once private conversations, and there are products like the Amazon Echo, which, like Apple's Siri, answers questions and plays music on demand, but also records your conversations to enhance voice recognition technology.[47] While television manufacturers, Amazon, and Apple claim that they do not use this information for marketing purposes, the growing capacity to collect and analyze one's most intimate conversations shows that historical boundaries for information-collection are under siege by evolving technology.[48]

This information—what some call our "digital exhaust"—gives data brokers and their customers the raw materials they need to use algorithms to categorize individuals into "segments" that are then sold to marketers looking to target specific slices of the population.[49] Some of these segments are based on obvious, exogenous characteristics and are not surprising; indeed, they appear to be commonsense ways in which to categorize discrete groups, such as college students, homeowners, automobile owners, dog owners, and golfing enthusiasts.[50]

But many of these segments raise serious questions, including, for example, those categorizing consumers who have modest financial means, including, for example, "Modest Wages," "Payday Loan Seekers," individuals who are "Financially Challenged," and "Thrifty Elders" (*i.e.*, individuals "in their late 60s and early 70s in 'one of the lowest income clusters'"), those that are race specific or appear to be proxies for race, including, for example, "African American Payday Loan Interest," "Urban Scramble" and "Mobile Mixers" (both of which "include a high concentration of Latino and African-American consumers with low incomes"), those who are ill, such as "Ailments and Pain Sufferers," "Diabetes Interest," and "Cholesterol Focus," or those who might engage in risky behavior, like "Biker/Hell's Angels."[51] And these categories are just the ones formulated and marketed by data brokers. Marketers often ask for specialized lists that target very discrete segments of individuals, ranging from people with bipolar disorder, specified sexual orientations, or those likely to inherit money soon or hold certain political or religious beliefs.[52]

---

47. Erica Fink & Laurie Segall, *Your TV Might be Watching You*, CNN MONEY (Aug. 1, 2013 11:32 AM), http://money.cnn.com/2013/08/01/technology/security/tv-hack/index.html; David Carnoy Crist, *Amazon Echo Review: The Smart Speaker that Can Control Your Whole House*, CNET (Mar. 3, 2016), http://www.cnet.com/products/amazon-echo-review/.

48. *See* INTERNET OF THINGS, *supra* note 44, at 1-2.

49. DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 19.

50. *Id.* at App. B: Illustrative List of Data Elements and Segments.

51. *Id.* at 20-25, App. B.

52. *See id.* at 25.

All of this raises a basic question: Should we worry that data brokers are collecting all of this information? The answer depends on exactly what the data is used for. Some uses seem unobjectionable. Marketers use this data to identify, with as much precision as possible, which groups of people, out of the more than 320 million Americans, are most likely to buy a particular product.[53] Data brokers argue that their ability to segment markets surgically provides an efficient marketplace that benefits us all.[54] And they have a point. These analytic tools give marketers access to exactly the audience they want to reach, and no more.[55] Targeted advertising saves marketers money by sparing them the expense of trying to sell products to disinterested consumers.[56] At the same time, targeted advertising saves consumers from the constant bombardment of ads for products in which they have no interest.[57] Everyone wins. Or at least that is what data brokers claim.[58]

But that is not the full story. Marketing based on big data may improve market efficiency, but it also poses three significant risks to consumers. Let us take a look at three of these risks, explain the threat they pose to consumers, and offer some preliminary thoughts about how society should go about mitigating these risks.

II. THE RISKS POSED BY BIG DATA AND ALGORITHMIC DECISION-MAKING

A. Ubiquitous Data Collection, Data Breach, and Identify Theft

As described above, the raw material of our data-driven economy is personally-identifiable information. To enable the transmission of data, our economy is highly networked so that data passes seamlessly from companies that collect massive amounts of personal data (*e.g.*, retailers, service providers, telecommunications companies, and insurers) to companies that aggregate that data, mainly data-brokers and analytic

53. *See id.*
54. *See* DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 3.
55. *Id.* at 3, 25.
56. *Id.* at 3.
57. *Id.*
58. The advertising industry's argument in support of targeting advertisements deliberately skips over a different, but nonetheless deeply problematic, risk. To engage in surgically targeted advertising, an advertiser needs to know a great deal about an individual's preferences—that is, the advertiser (or its partners) needs to have developed an especially rich and detailed profile of the individual. To be sure, that profile can be used to discern preference so that the individual receives advertisements for watches rather than laundry soap. But that profile can also be used to adapt (the industry calls this process "curating") the content an individual receives, including news feeds, political messages, and other non-advertising content. The ability to shape the content an individual receives may, in turn, influence or shape the individual's preferences, and that too may be a goal of certain advertisers.

companies, who then resell it.[59] One consequence of ubiquitous data collection is that the data-acquiring entities store massive amounts of personal information, and thus become "honey pots," or targets of opportunity, for malicious hackers intending to steal that data.[60] Unfortunately, too often these companies are poor data stewards and fail to implement data security measures commensurate with the sensitivity and volume of the information they store.[61] They are thus easy prey for hackers.

For this reason, the most urgent threat to consumer privacy, and the most disruptive aspect of the new "big data" economy, is that sensitive data routinely ends up in the hands of identity thieves.[62] Most prized are Social Security numbers, credit card information, passwords or authentication credentials, and information that might facilitate the theft of an individual's medical insurance.[63] Hackers steal this information and then auction it off to identity thieves in secret, or "dark," Internet sites.[64] Once they acquire the data, the thieves then use, and often reuse, it for one or more forms of identity theft.

Make no mistake, data breaches are responsible for the meteoric rise of identity theft in the United States, and they are the predictable debris of an Internet economy that places too little value on data security. The Federal Trade Commission ("FTC") has long collected complaints from individuals whose identity has been stolen and tries to assist these individuals in reclaiming their own identity. In 2014, over 330,000 individuals filed identity theft complaints with the FTC, and these complaints are just the tip of a much larger iceberg.[65] According to the Department of Justice, "an estimated 17.6 million persons, or about 7 percent of U.S. residents age 16 or older, were victims of at least one incident of identity theft in 2014."[66]

The rate of identity theft has risen every year since 2000 in virtual lockstep with the rate of data breaches.[67] Of course, it is rarely possible to link an act of identity theft to a specific data breach, but the data stolen in

59. *See id.* at 11.

60. *See* DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 48-49.

61. *See id.* at 55.

62. *See id.*

63. *See id.* More broadly, these thieves covet consumer profiles that provide information that might enable them hack into accounts by predicting passwords, helping to answer "challenge questions" (e.g., "what is your mother's maiden name?"), or providing other authentication credentials. *Id.*

64. *Computer Hacking and Identity Theft*, PRIVACY MATTERS, http://www.privacymatters.com/identity-theft-information/identity-theft-computer-hacking.aspx (last visited Mar. 14, 2016).

65. FED. TRADE COMM'N, CONSUMER SENTINEL NETWORK DATA BOOK FOR JANUARY-DECEMBER 2014 5 (Feb. 2015), https://www.ftc.gov/system/files/documents/reports/consumer-sentinel-network-data-book-january-december-2014/sentinel-cy2014-1.pdf [hereinafter CONSUMER SENTINEL NETWORK DATA BOOK].

66. Press Release, Bureau of Justice Statistics, 17.6 Million U.S. Residents Experienced Identity Theft in 2014 (Sept. 27, 2015), http://www.bjs.gov/content/pub/press/vit14pr.cfm.

67. CONSUMER SENTINEL NETWORK DATA BOOK, *supra* note 65, at 5.

breaches provides the tools for identify theft. And identity theft is big business: According to the most recent Department of Justice statistics, in 2012, identity theft cost the U.S. economy $24 billion dollars, $10 billion more than all of the losses attributable to property crimes during the same time period.[68]

Virtually no one is spared from the adverse effects of data breaches. According to the Financial Services Roundtable, the association representing the nation's biggest banks, approximately 110 million Americans had their data exposed during 2014.[69] And the pace of data breach is not slowing. It is thus not unreasonable to assume that most Americans have, at some point in the past few years, had sensitive information about them compromised by a data breach.

Identity thieves can wreak havoc on the lives of those whose identity they assume.[70] They make unauthorized credit card charges. They use health insurance information to obtain expensive medical care, and by so doing, often trigger the denial of medical care to the insured. They use Social Security numbers to file fake tax returns, to obtain other government benefits, and, worst of all, to assume the identities of children. Unfortunately, victims of identity theft have few avenues of redress.

In cases involving financial data, remediation (but not compensation) may be possible, but it places significant burdens on the individual whose data has been compromised. As to financial loss, "Regulation Z," permits credit card holders to insist that the bank that issued the holder's card "charge back" unauthorized charges.[71] In theory, this charge back right protects a consumer from fraudulent charges. But that theory depends on two assumptions: 1) that the consumer is able to identify all or most unauthorized charges, and 2) that the consumer receives timely notification of the risk of unauthorized charges. Neither assumption is sound.

For example, the global hotel chain Wyndham experienced three data breaches within an eighteen month span.[72] There is good reason to believe that the credit card files of over 600,000 Wyndham customers ended up in the hands of Russian organized crime.[73] Presumably, the credit card files

---

68. ERIKA HARRELL & LYNN LANGTON, VICTIMS OF IDENTITY THEFT, 2012 6 (Dec. 2013), http://www.bjs.gov/content/pub/pdf/vit12.pdf.

69. Erin Kelly, *Officials Warn 500 Million Financial Records Hacked*, USA TODAY, (Oct. 20, 2014), http://www.usatoday.com/story/news/politics/2014/10/20/secret-service-fbi-hack-cybersecuurity/17615029/.

70. *Computer Hacking and Identity Theft*, *supra* note 64.

71. 12 C.F.R. 226.12(c). Reg. Z does not apply to debit cards, but some banks do permit charge backs from debit cards for clearly unauthorized transactions. 12 C.F.R. § 226.5a(a)(5).

72. FTC v. Wyndham Worldwide Corp., 799 F.3d 236, 241-42 (3rd Cir. 2015); Complaint for Injunctive and Other Equitable Relief at 13-17, FTC v. Wyndham Worldwide Corp., 10 F.Supp.3d 602 (D.N.J. 2014) (No. 13-1887).

73. *Wyndham*, 799 F.3d at 242; Complaint, *supra* note 71, at 17.

stolen from Wyndham were sold to criminals who know how to sneak charges onto credit card statements (the Russian Mafia did not steal the information just for fun). Fake charges do not appear on consumers' credit card statements as "Russian Mafia;" rather, they are carefully disguised to look like actual charges. Complicating matters, consumers were notified of the Wyndham breaches a year or more after they occurred.[74] It is hardly reasonable to expect consumers to have all of their statements available and for them to be able to determine what charges were unauthorized months or even years after-the-fact.

To be sure, there are remedial measures a consumer can take to prevent future financial loss flowing from a breach.[75] But doing so is a daunting task. The FTC gives excellent advice and provides consumers with tools to report the theft to the police, creditors, and others.[76] But carrying out these measures takes time, requires access to all of one's financial information, and generally requires access to, and the ability to use, an Internet-enabled computer. To avoid future financial losses, consumers can place fraud and extended fraud alerts by contacting each of the three credit bureaus (Equifax, Experian, and Trans Union).[77] Consumers can also place credit freezes, which bar access to new credit by anyone claiming the individual's identity, again by arranging these freezes with the credit bureaus, although here consumers may have to pay a fee for placing, lifting, and removing the freeze.[78] For consumers in the market for credit, placing a credit freeze can be problematic.[79] But these are measures that can and do provide some measure of protection against future losses.[80] Again, they do nothing to make consumers whole.

But at least victims of financial identity theft have some recourse. That is not true for those victimized by more virulent, and increasingly more common, forms of identity theft. There are no off-the-shelf remedial measures when personal data is used to engage in taxpayer identity theft, in medical identity theft, or in children's identity theft. To be sure, the IRS has a process that taxpayers can follow when they have been victimized.[81] But

74. *Wyndham*, 799 F.3d at 242; *see* Complaint, *supra* note 71, at 16-17.

75. *Computer Hacking and Identity Theft*, *supra* note 63 (listing measures consumers can take to protect themselves from identity theft).

76. Fed. Trade Comm'n, *Identity Theft Recovery Steps*, IDENTITYTHEFT.GOV, https://identity theft.gov/ Steps (last visited Mar. 14, 2016).

77. *Id.* ("What to Do Right Away").

78. Fed. Trade Comm'n, *Credit Freeze FAQs*, CONSUMER.FTC.GOV, https://www.consumer. ftc.gov/articles/0497-credit-freeze-faqs (last visited Mar. 14, 2016).

79. *Id.*

80. *Id.*

81. Internal Revenue Service, *Taxpayer Guide to Identity Theft*, IRS.GOV, https:// www.irs.gov/uac/Taxpayer-Guide-to-Identity-Theft (last visited Mar. 14, 2016). The FTC also provides advice for tax identity theft. FED. TRADE COMM'N, CONSUMER INFORMATION: TAX-RELATED IDENTITY

taxpayers complain bitterly about the IRS's response to claims of taxpayer identity theft.[82] For medical identity theft, there is no set process to follow.[83] The onus falls on the victim to wage a campaign, provider-by-provider, to reclaim his or her own identity.[84] And for child identity theft, the problem is that, generally, the theft is not discovered until the child reaches age eighteen, and at that point, unwinding the damage done by the theft can be a monumental task.[85]

There are, unfortunately, revelations of personal data that are not capable of remediation.[86] For example, no amount of credit monitoring is going to console the men and woman who signed up with Ashley Madison, a web site for married people seeking extra-marital relationships, and have seen their marriages dissolve in the wake of a data breach.[87] No credit freeze is going to comfort those who were prescribed Prozac and had their identities revealed by Eli Lilly.[88] And there is no peace for consumers who rented computers that had DesignerWare Software installed, which was used surreptitiously to take photographs of them and the families in their homes without their knowledge.[89] These photographs included pictures of children, household visitors, individuals not fully clothed, and couples engaged in intimate activities.[90] Again, there is no "remediation" for these outrageous assaults on privacy.

THEFT, https://www.consumer.ftc.gov/articles/0008-tax-related-identity-theft (last visited Mar. 14, 2016).

82. *See Tax Refund ID Theft is Growing "Epidemic": U.S. IRS Watchdog*, RUETERS (Nov. 7, 2013), http://www.reuters.com/article/us-usa-tax-refund-idUSBRE9A61HB20131107; Susan Tompor, *IRS to Warn 685,000 Tax Filers of ID Theft*, DETROIT FREE PRESS (Feb. 26, 2016), http://www.freep.com/story/money/personal-finance/susan-tompor/2016/02/26/get-transcript-trouble-builds-irs-scammers/80995582/.

83. *See* FED. TRADE COMM'N, MEDICAL IDENTITY THEFT, https://www.consumer.ftc.gov/articles/0171-medical-identity-theft (last visited Mar. 14, 2016).

84. Laura Shin, *Why Medical Identity Theft is Rising and How to Protect Yourself*, FORBES (May 29, 2015), http://www.forbes.com/sites/laurashin/2015/05/29/why-medical-identity-theft-is-rising-and-how-to-protect-yourself/#44451905e200.

85. *See* FED. TRADE COMM'N, CHILD IDENTITY THEFT, https://www.consumer.ftc.gov/articles/0040-child-identity-theft (last visited Mar. 14, 2016); *ITRC Fact Sheet 120: Identity Theft and Children*, IDENTITY THEFT RESOURCE CENTER, http://www.idtheftcenter.org/Fact-Sheets/fs-120.html (last visited Mar. 14, 2016).

86. Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1705 (2010).

87. *See* Kim Zetter, *Hackers Finally Post Stolen Ashley Madison Data*, WIRED (Aug. 18, 2015), http://www.wired.com/2015/08/happened-hackers-posted-stolen-ashley-madison-data/.

88. Press Release, Fed. Trade Comm'n, Eli Lilly Settles FTC Charges Concerning Security Breach (Jan. 18, 2002), https://www.ftc.gov/news-events/press-releases/2002/01/eli-lilly-settles-ftc-charges-concerning-security-breach.

89. Richard Larner, *DesignerWare Software: Companies Agree to Stop Snooping on People's Home Computers*, HUFFINGTON POST (Sept. 26, 2012), http://www.huffingtonpost.com/2012/09/26/designerware-settlement_n_1917255.html.

90. *Id.*

How do consumers protect themselves in an era of big data? There are no easy answers. So far, the FTC has brought nearly sixty enforcement actions against companies that failed to have reasonable data security measures in place.[91] But these actions have not had an observable deterrent effect, no doubt because the remedies available to the Commission are limited to forward-looking injunctive relief.[92] The Commission has no authority to impose civil penalties or other monetary relief; it can only force companies to upgrade their systems and conduct audits to ensure that their systems provide reasonably adequate security going forward.[93] For these reasons, the Commission has long urged Congress to provide the agency with much stronger remedies, including civil penalties.[94]

Congress needs to recognize that cyber-security is a top priority in our increasingly networked economy. Not only do data breaches exact a serious toll on our economy, but at present, all of the evidence underscores that there are insufficient economic incentives to push companies to have reasonable security measures.[95] Sasha Romanosky, a researcher with the Rand Institute, recently conducted a comprehensive study of publicly-disclosed data breaches to see the impact the breaches had on the companies that experienced a breach.[96] He concluded that the economic loss to the companies was at most modest and transitory, and that the risks were inadequate to push the companies to upgrade their data security.[97] Most costs are externalized and are borne by consumers, banks, and affiliates.[98] The conclusion was sobering—having lax security measures and risking, and even experiencing, a breach was a reasonable approach for data-collecting companies as a matter of economics.[99]

One reason why the economic incentives cut in the wrong direction is that courts have been reluctant to recognize the harms to consumers from

---

91. *Privacy & Data Security Update (2015)*, FED. TRADE COMM'N, https://www.ftc.gov/reports/privacy-data-security-update-2015 (last visited Mar. 25, 2016); *see* Solove & Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 600 (2014) ("[T]he number of FTC enforcement actions has not been particularly voluminous.").

92. Solove & Hartzog, *supra* note 91, at 599.

93. *See id.*

94. Press Release, Fed. Trade Comm'n, FTC Testifies on Efforts to Protect Consumers of Financial Services; Urges New Tools for Stronger Enforcement Authority (Mar. 24, 2009), https://www.ftc.gov/news-events/press-releases/2009/03/ftc-testifies-efforts-protect-consumers-financial-services-urges.

95. *See id.*

96. Sasha Romanosky, *Examining the Costs and Causes of Cyber Incidents* (Fed. Trade Comm'n PrivacyCon, Jan. 14, 2016), . , https://www.ftc.gov/system/files/documents/public_comments/2015/10/00027-97671.pdf (last visited Mar. 29, 2016).

97. *See id.* at 20-21.

98. *See id.*

99. *See id.*

data breaches, thereby insulating these companies from any form of consumer redress.[100] To be sure, class action lawsuits are routinely filed in the wake of most breaches, but they are generally dismissed because the court concludes that consumers have failed to allege with adequate particularity that they have suffered harm that can be attributable to a specific breach.[101] These decisions often give short shrift to the reality that pinpointing identity theft to a specific breach is often a fool's errand, and that, as a consequence of the breach (consider the massive Target breach), consumers generally have to spend a good deal of time getting new credit cards, arranging for recurring charges to be moved to newly-issued cards, and vigilantly monitoring their card statements for months or years after the breach.[102] Consumers should be entitled to redress for at least these injuries, if not more.

B. Unrestrained Collection of Sensitive, Personal Data

One way to mitigate the harms that result from data breaches is to reduce the amount of personal information that is collected in the first place. There are laws that bar the collection of certain categories of data, and a major issue in the debate over big data is whether collection restrictions make sense these days.[103] Big data proponents argue that the traditional requirements of notice and choice for data collection are antiquated, and cannot reasonably be applied where data is increasingly collected through sensors and machines that have no interface with consumers.[104] They also point out that data collected for one purpose can at times be highly valuable for uses that were not anticipated when the data was first collected.[105] For instance, health information that may be used to treat a patient (who gave informed consent to the treatment) may provide important insights in designing new and better therapies. Their answer is to ease restrictions on collection and to deter data-use practices that society deems inappropriate by tailored use restrictions.[106]

---

100. *See* Sasha Romanosky, *Empirical Analysis of Data Breach Litigation*, 11 J EMPIRICAL LEGAL STUD. 74, 100 (2014).

101. *See id.* at 102; *See generally Litigation*, DATA BREACH TODAY, http://www.databreachtoday.com/litigation-c-320 (last visited Mar. 25, 2016) (collecting cases).

102. *See Identity Theft Recovery Steps*, *supra* note 76.

103. *See* Bedoya & Vladeck, *supra* note 37.

104. *See id.*

105. *See* DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 3.

106. WORLD ECONOMIC FORUM, BIG DATA, BIG IMPACT: NEW POSSIBILITIES FOR INTERNATIONAL DEVELOPMENT 2 (2012). Where truly sensitive data is concerned, use restrictions provide little if any meaningful protection to consumers. For one thing, use restrictions are generally imposed only after serious abuses arise. For another, violations of use restrictions are punishable only in ways that are intended deter future misuse, but do little or nothing to remediate the harm to consumers. For yet another, use restrictions are no help in securing sensitive data from data breaches. And finally,

This argument is, unsurprisingly, backed by marketers. After all, highly personal information is the coin of the realm for digital advertisers. Ubiquitous data collection is key. Digital marketing depends on the steady flow of personal information into analytic companies.[107] But data collection of the scope required to support highly-personalized marketing is not inevitable for several reasons.

The first reason is that as consumers become aware of the ubiquity of data collection, they are starting to push back, and regulators and some companies in the business of privacy enhancement are seeking to restore to consumers control over their own data.[108] Since 2010, the FTC has pushed for a "Do Not Track" option to provide consumers significant control over the collection of personal data while browsing the internet.[109] Although efforts to negotiate a Do Not Track regime fell apart, the FTC's effort has nonetheless borne fruit. All of the major browsers (Microsoft's Internet Explorer, Mozilla's Firefox, and Apple's Safari) offer Do Not Track options that permit anonymous browsing, and in 2012 Microsoft issued its Internet Explorer 10, which by default is set to Do Not Track.[110] To be sure, advertisers have resisted compliance.[111] But that resistance may be starting to weaken. Mozilla and Apple are considering setting their browsers to a default Do Not Track setting as well.[112] The advertising industry now projects that up to half of browser activity will soon be sending out Do Not Track signals, a development which has prompted the industry to reconsider its position.[113] Although it is too early to declare victory for Do Not Track, it is also too late to deny that consumers are increasingly demanding control

---

use restrictions are subject to attack as a species of viewpoint discrimination, which raises thorny First Amendment issues. *See generally* Sorrell v. IMS Health Inc., 131 S. Ct. 2653 (2011).

107. *See* DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 3.

108. *See* Bedoya & Vladeck, *supra* note 37.

109. FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: A PROPOSED FRAMEWORK FOR BUSINESS AND POLICYMAKERS, PRELIMINARY FTC STAFF REPORT 63-66 (2010) [hereinafter PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE].

110. Craig Buckler, *Microsoft IE10 and Its "Do Not Track" Default Settings*, SITEPOINT (Oct. 19, 2012), http://www.sitepoint.com/ie10-do-not-track/; *Do Not Track FAQ*, MOZILLA, https://www.mozilla.org/en-US/dnt/; *Safari 9 (El Capitan): Ask Websites Not to Track You*, APPLE (Sept. 30, 2015), https://support.apple.com/kb/PH21416?viewlocale=en_US&locale=en_US.

111. Elise Ackerman, *Google and Facebook Ignore "Do Not Track" Requests, Claim They Confuse Consumers*, FORBES (Feb. 27, 2013), http://www.forbes.com/sites/eliseackerman/2013/02/27/big-internet-companies-struggle-over-proper-response-to-consumers-do-not-track-requests/; Elizabeth Dwoskin, *Yahoo Won't Honor "Do Not Track" Requests from Users*, WALL ST. J. DIGITS (May 2, 2014), http://blogs.wsj.com/digits/2014/05/02/yahoo-wont-honor-do-not-track-requests-from-users/.

112. Jim Edwards, *Death of the Cookie: How the Web's All-Seeing Tracking Device Could Meet Its End*, BUS. INSIDER (May 1, 2013), http://www.businessinsider.com/death-of-cookies-2013-4.

113. Jim Edwards, *Advertisers May Capitulate on Use of Tracking Cookies: "That Is No Longer Tenable"*, BUS. INSIDER (July 9, 2013). http://www.businessinsider.com/ad-business-admits-do-not-track-cookies-have-won-2013-7.

over tracking and that the market is responding with more and better tools consumers can use to protect their data.[114]

Growing consumer awareness may also spur legislation that either restricts collection of certain categories of data or places strict limitations on their use. After all, it was Justice Scalia—no defender of privacy rights—who warned that society needs to guard against the "power of technology to shrink the realm of guaranteed privacy."[115] There is ample precedent for legislation that limits data collection, subject, of course, to explicit consent by the consumer. For instance, the Children's Online Privacy Protection Act forbids the tracking of children ages twelve and under on the internet without express parental permission.[116] The Wiretap Act forbids the use of electronic means to record a conversation without consent.[117] Many state laws forbid the recording of telephone conversations without the consent of both parties.[118] And the recently-enacted Driver Privacy Act protects the privacy of the event data recorder now common in many cars (like the "black box" on commercial airplanes) by explicitly limiting access to it without the driver's permission, absent special circumstances.[119]

There are certainly other categories of information that perhaps should also be off-limits to data collectors without affirmative, express consent, like geolocation data, information about family and health matters, and other highly sensitive personal information. The FTC has long taken this position, and has urged Congress to enact baseline privacy legislation that would restore control over sensitive data to individuals, and not corporate data-collectors.[120]

Another avenue for policy-makers to explore is borrowing from our Fourth Amendment jurisprudence to consider using the limits on government surveillance as a guide to restricting commercial surveillance. The key Fourth Amendment case remains *Katz v. United States*,[121] which

---

114. *See* Simon Davies, *Three Healthy Indications That Online Privacy May Have Turned a Corner*, PRIVACY SURGEON (July 11, 2013), http://www.privacysurgeon.org/blog/incision/three-healthy-indicators-that-online-privacy-may-have-turned-a-corner/; Mark Little, *"Little Data": Big Data's New Battleground*, OVUM (Jan. 29, 2013), http://www.ovum.com/little-data-big-datas-new-battleground/. The ability of tracking companies to track consumers across devices adds to the concern about ubiquitous data collection of personally identifiable information. *See generally* FED. TRADE COMM'N, CROSS-DEVICE TRACKING (Nov. 16, 2015), https://www.ftc.gov/news-events/events-calendar/2015/11/cross-device-tracking.

115. Kyllo v. United States, 533 U.S. 27, 34, 40 (2001).

116. 15 U.S.C. §§ 6501(1), 6502(A)(1) (2016).

117. 18 U.S.C. § 2511(1)(b) (2016).

118. *Summary of Consent Requirements for Taping Telephone Conversations*, AAPS, http://www.aapsonline.org/judicial/telephone.htm (last visited Mar. 14, 2016).

119. Driver Privacy Act of 2015, Pub. L. No. 114-94, § 24301, 129 Stat. 1312 (2015).

120. *See* PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE, *supra* note 109, at 63-66

121. 389 U.S. 347 (1967).

established, albeit through Justice Harlan's concurring opinion, the principle that reasonable expectations of privacy are the Fourth Amendment's touchstone and must be safeguarded by society.[122] According to Justice Harlan, there are two facets of privacy expectations; one is the individual's subjective expectation, but the second, and often more important (and contestable) consideration is whether there is an objective, reasonable expectation that a communicative act will remain private.[123]

What is important here is that Justice Harlan did not pick "reasonableness" out of thin air. The Court had applied reasonableness considerations in prior Fourth Amendment cases, and did so for the simple reason that reasonableness is a standard that pervades Anglo-American law.[124] So much of our foundational law—tort, contract, property—is based on our elusive search for the "reasonable person" or to determine what constitutes "reasonable" conduct.

Where would such an inquiry take us? The Court's Fourth Amendment decisions guide us in ways that reflect at least widely accepted societal norms, if not consensus, about where to draw the line between appropriate acts of data-gathering and inappropriate surveillance.[125] And it is fair to ask why society should permit corporate surveillance in those areas where the Supreme Court has found that the Fourth Amendment places strict limits on warrantless government surveillance. Take smartphone tracking. The Court's recent decision in *United States v. Jones*[126] held that law enforcement officers violated the defendant's Fourth Amendment rights by putting a GPS device on the defendant's car and using it for nearly a month to track the defendant's whereabouts.[127] The crux of the Court's ruling is that the constant monitoring of the defendant's movements, without a warrant, violated the defendant's reasonable expectation of privacy.[128] It is hardly a stretch to suggest that it may not be "reasonable" for mobile apps to collect and disseminate exactly the same geolocation tracking information without the consumer's meaningful, affirmative consent.

The same can be said of the Court's ruling in *Riley v. California*,[129] holding that the government may not rummage through a suspect's

122. *Id.* at 361 (Harlan, J. concurring).
123. *Id.*
124. *See id.* at 361-62.
125. *See id.*
126. 132 S. Ct. 945 (2012).
127. *Id.* at 949. Justice Sotomayor, in her concurring opinion, emphasized that "GPS monitoring generates a precise, comprehensive record of a person's public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations." *Id.* at 955 (Sotomayor, J. concurring).
128. *Id.* at 950.
129. 134 S. Ct. 2473 (2014).

smartphone without a warrant because doing so violates the defendant's reasonable expectation that information stored on his smartphone's hard drive, including pictures, cell phone contacts, texts messages, and video clips, would be private.[130] Surely the Court's ruling in *Riley* suggests that mobile apps should not be able to pull down contact information, texts, pictures, videos, or other information stored on a smartphone's hard drive without the consumer's clear, affirmative consent.

As a final example, consider the Supreme Court's ruling in *Kyllo v. United States*,[131] holding that the government could not use a thermal-imaging device to detect heat sources within a private home (an indication of marijuana production) without a warrant.[132] Now consider Google's Nest and competing devices, which do not simply *detect* heat sources, but will actually *control* the heat sources in homes. Following *Kyllo,* is it fair to argue that placing devices like sensors inside the home to collect and share information with third parties about how the homeowners live their lives requires heightened consent?

Just because technology has the power "to shrink the realm of guaranteed privacy"[133] does not mean that society has to stand idly by as technology erodes our private spheres. As consumers become more aware of the extent to which their personal data is being harvested by commercial entities, the more wary they become. It is time for policy-makers to engage on these issues and put control over personal data back in the hands of individuals who, provided that they are fully informed, can make their own decisions about data collection and data use.[134]

---

130. *Id.* at 2489, 95.

131. 533 U.S. 27 (2001).

132. *Id.* at 40.

133. *Id.* at 34.

134. There is one other important Fourth Amendment case that oddly has had little resonance in civil privacy cases. *Bivens v. Six Unnamed Agents of Federal Bureau of Narcotics*, 403 U.S. 388 (1971), gave rise not only to an implied right of action to enforce Fourth Amendment rights but also to a right to money damages to provide a remedy for Fourth Amendment deprivations. Bivens' privacy and liberty were invaded when federal agents, "acting under claim of federal authority, entered his apartment and arrested him for alleged narcotics violations. The agents manacled petitioner in front of his wife and children, and threatened to arrest the entire family. They searched the apartment from stem to stern. Thereafter, petitioner was taken to the federal courthouse in Brooklyn, where he was interrogated, booked, and subjected to a visual strip search." *Id.* at 389. Bivens' injury was, at its core, an assault on his dignity; an injury perhaps made graver because it was inflicted by government agents rather than private parties, but not different in kind. Writing for the Court, Justice Brennan held that damages are appropriate both to remedy Bivens' injury and to deter federal officials from future Fourth Amendment violations. *Id.* at 95-96. Justice Harlan's more detailed concurrence offers the same justifications, but drives home the futility of remedies other than damages with his memorable admonition: "For people in Bivens' shoes, it is damages or nothing." *Id.* at 410 (Harlan, J., concurring). *See also* Monroe v. Pape, 365 U.S. 473 (1961) (recognizing that the Civil Rights Act, 42 U.S.C. §1983, provides a right to bring a damage action against state and local officials who engage in warrantless searches and seizures). It is time for courts handling civil privacy cases to come to the realization that in privacy cases, the remedy is "damages or nothing," and that "nothing" is not an attractive option. It not only leaves injured parties to

C. The Use of Algorithmic Decision-Making to Make Consequential Decisions

Staunching the flow of personal information into the repositories of data brokers will not, by itself, solve the problems with algorithmic decision-making.[135] That requires other measures. As noted earlier, the challenge policy-makers face is how to leverage the promise of machine learning algorithms to make better decisions, while ensuring that the correlations they draw are not arbitrary and do not discriminate against individuals on any bases forbidden by law.[136]

Marketing data is often fed into algorithms to determine whether individuals qualify for a product or service, and if so, the terms the individual should be offered.[137] These determinations can have serious ramifications. For one thing, they may affect the choices consumers are given, reinforcing existing disparities.[138] Suppose a financial institution is selecting consumers to target with offers of prime, low-interest loans. And suppose that some deserving consumers are wrongly characterized as credit risks, not based on their own credit history, but because others who may share some characteristics with them are predicted to be poor credit risks.[139] Those wrongly categorized consumers are not just denied the immediate benefit of receiving a favorable credit offer, but, making matters worse, they are also never informed of the mis-categorization. As a result, they cannot take steps to ensure that the error does not recur or is not compounded in the future.[140]

Increasingly, prices are variable, and the price a consumer is charged may depend on what an algorithm determines to be the optimal price.[141] For example, "online companies may charge consumers in different zip codes different prices for standard office products."[142] These pricing

---

bear the loss they did not cause, but it also removes a critical deterrent, thereby inviting the careless handling of private information.

135. *See* Bedoya & Vladeck, *supra* note 37.

136. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 2; *see* Exec. Office of the President, Big Data: Seizing Opportunities, Preserving Values (May 2014), https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.

137. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 2.

138. *Id.*

139. *Id.* at 9-10.

140. PERSIS YU, JILLIAN MCLAUGHLIN & MARINA LEVY, BIG DATA: A BIG DISAPPOINTMENT FOR SCORING CONSUMER CREDIT RISK, NAT'L CONSUMER LAW CTR. 10 (2014), http://www.nclc.org/images/pdf/pr-reports/report-big-data.pdf.

141. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 11; *see* David Streitfeld, *It's Discounted, but Is It a Deal? How List Prices Lost Their Meaning*, N.Y. TIMES, (Mar. 6, 2016), http://www.nytimes.com/2016/03/06/technology/its-discounted-but-is-it-a-deal-how-list-prices-lost-their-meaning.html?_r=0.

142. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 11.

determinations may result in "consumers in poorer neighborhoods having to pay more for online products than consumers in affluent communities, where there is more competition from brick-and-mortar stores."[143] In this way, consumers living in these poorer communities would be the losers, not winners, in the competitive online marketplace.[144]

Finally, algorithmic decision-making may give companies new ways to justify disparate treatment of groups based on correlations.[145] Suppose an algorithm concludes that people who write with block letters, or need to load new programs onto their computers to complete certain tasks, frequently change jobs. Should that correlation be used to limit their employment opportunities? And if these correlations were used by an employer, would any job applicant ever realize that he or she was not hired for these reasons?[146]

And that's the problem. The process of engaging in algorithmic decision-making is one that is opaque and complex, and can effectively mask discrimination.[147] Under applicable laws, discrimination can generally be shown in two ways. One is with evidence of "disparate treatment," that is, proof that a company deliberately fails to offer goods or services to, for example, minorities or women, on the same terms it offers to others.[148] Those cases are rare. More often, discrimination is far more subtle and is shown through discriminatory impact theories, that is, with evidence that the challenged practices have a disproportionately adverse effect on minorities or other protected groups and are otherwise unjustified by a legitimate rationale.[149]

But regulators will have a hard time uncovering discrimination when the decision is made by machine-learning algorithms that process mounds of data.[150] To be sure, regulators can review the inputs to see whether any of the data used involves forbidden categories or categories that may be proxies for forbidden categories. For instance, if a key input is an individual's "zip code plus four"—that is, the highly specific zip codes that often are highly correlated with race, wealth, home ownership, political orientation, and more—that may be a warning sign that something is amiss. Generally, however, algorithms are fed dozens if not hundreds of pieces of

143. *Id.*
144. *See id.*
145. *Id.* at 10-11.
146. Some of the practices mentioned below might, at least arguably, violate several existing statutes, including the Fair Credit Reporting Act and the Equal Credit Opportunity Act. *Id.* at 12-23.
147. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 10-11.
148. *Id.* at 18-19.
149. *Id.*; *see e.g.*, Texas Dept. of Hous. and Cmty. Affairs v. The Inclusive Cmty. Project, Inc., 135 S. Ct. 2507, 2513 (2015).
150. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 8.

information to process, and even sifting through the inputs may be a challenge.[151]

Assuming that regulators would even have the technical expertise to do so, they would rarely, if ever, be able to detect telltale signs of discriminatory intent or impact while looking at the black box of the algorithm. Machine learning algorithms "learn" as they process data, so the process is fluid, not static, and the factors the algorithm deems significant change over time.[152]

For these reasons, regulators are left with few tools to ensure that algorithmic decision-making does not result in discrimination.[153] One approach is the one taken by the Federal Reserve Board, which sets regulatory requirements for the Equal Credit Opportunity Act.[154] Under "Regulation B," lenders who use non-traditional factors in making credit determinations must implement "empirically derived, demonstrably and statistically sound, credit scoring system[s]."[155] These tools must be "developed and validated using accepted statistical principles and methodology," and must be subject to ongoing evaluation and review.[156] The other approach is to require the companies that engage in algorithmic decision-making to conduct after-the-fact audits to determine whether the use of the algorithm has a discriminatory effect.[157] Compelling companies to undertake these audits is, at least for now, beyond the power of most of the federal agencies that oversee lenders.[158] To be sure, the agencies themselves might conduct such an audit, but doing so would pose an enormous challenge to regulators, who are under-resourced and may not have the expertise required to design and carry out a sufficiently robust audit.

---

151. DATA BROKERS: A CALL FOR TRANSPARENCY, *supra* note 30, at 31. As the FTC's Data Broker report makes clear, data in the hands of data brokers is rife with inaccuracies. These inaccuracies make little difference when targeting consumer for ads for soap or soup, but they can have serious consequences when the products are expensive (and may result in disadvantageous differential pricing) or when the offers depend on judgments of creditworthiness. *See id.* at 36-37 (data brokers tend to implement more rigorous measures to ensure data quality for identity verification products than for marketing products).

152. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 28; *see generally* Harry Surden, *Machine Learning and the Law*, 89 Wash. L. Rev. 87, 89-95 (2014).

153. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 28.

154. 15 U.S.C. § 1691 *et seq.* (as amended by the Dodd-Frank Act).

155. *See* 12 C.F.R. § 1002.6(b). (2016) (requiring a sound, empirical credit scoring system when considering age as a factor).

156. 12 C.F.R. § 202.2 (2016).

157. BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?, *supra* note 2, at 32.

158. *See id.* at 19-21.

III. CONCLUSION

The era of big data analytics holds considerable promise for society. But as history has shown, the introduction of disruptive technologies comes at a cost. One cost of big data is its insatiable appetite for more data. As a result, data collection has become pervasive. Even our homes are under siege by would-be data collectors. Maintaining one's privacy no longer consists of locking one's front door. In today's networked world, maintaining privacy, if at all possible, is a full-time job.

But our big data world is still in its infancy. Google, Facebook, Twitter and the other Internet giants have yet to reach their teens. Smartphones have been with us for less than a decade. And we are still struggling to understand the impact these new technologies are having, and will have, on our society.

It is too early to fully measure big data's impact on consumer protection, but we already can see telltale signs of serious harm. Identity theft is now rampant. Invasions of privacy are everyday occurrences. And big data analytics pose risks of overt discrimination as sellers set prices based on who they think we are, and hard to root out disparate treatment based on impermissible factors like race, gender, sexual orientation, and age.

None of these harms is inevitable. There are means to combat them all. But society will have to decide how to protect consumers from these and other threats while, at the same time, harnessing the power of big data to do good. It is that challenge we have to meet.