Written Testimony of Duane Pozza
Partner, Wiley Rein LLP

Before the United States House Committee on Financial Services
Task Force on Financial Technology

Hearing entitled "Banking on Your Data: the Role of Big Data in Financial Services"

Thursday, November 21, 2019
Rayburn House Office Building, Room 2128

Chairman Lynch, Ranking Member Emmer, and Members of the Task Force on Financial Technology, thank you for the opportunity to appear today to discuss the role of big data in financial services.

I am a partner at Wiley Rein LLP, where my practice includes advising companies on the legal and regulatory framework for collecting, using, and managing consumer data, including in financial services. This includes counseling on U.S. and global data privacy laws, financial services laws and regulations, and emerging regulatory approaches and expectations around the use of artificial intelligence (AI) and machine learning technologies, which depend on large and sophisticated data sets. I previously worked at the Federal Trade Commission, including as an Assistant Director in the Division of Financial Practices in the Bureau of Consumer Protection. I helped organize the FTC FinTech Forum Series, which examined, among other things, the role of big data in financial services, including through a 2017 event that focused on the consumer-focused uses of artificial intelligence technology.[1]

Data-driven financial services hold enormous potential to improve consumers' financial lives. Companies can use consumer data responsibly to expand access to credit, provide customized financial advice, detect and prevent fraudulent behavior, and provide financial services at a lower cost. Companies are already using large and robust data sets to accomplish these objectives, and the development of machine learning and AI technologies will further advance what technology innovators can accomplish.

Companies using consumer data in innovative ways for financial decisions operate in an area with many significant laws and regulations on the books and multiple regulatory authorities. Companies must comply with well-established financial services laws, many of which implicate use of consumer data, as well as FTC guidance on data privacy and security. They must also comply, to varying degrees, with consumer privacy laws that reach across sectors, both on the international level (for example, the European Union's General Data Protection Regulation) and state level (for example, the California Consumer Privacy Act). State laws in particular threaten to create a piecemeal compliance framework and burden businesses that already have substantial

[1] *See FinTech Forum: Artificial Intelligence and Blockchain*, FTC EVENTS CALENDAR (Mar. 9, 2017), https://www.ftc.gov/news-events/events-calendar/2017/03/fintech-forum-blockchain-artificial-intelligence.

compliance obligations. The experience with California's law illustrates some of the challenges that companies face. As consumer data is increasingly used to provide better financial services, it is important to carefully consider consumer expectations and preferences around use of their information, and weigh the benefits that better financial services can bring and the significant cost of added regulation.

Using Consumer Data to Improve Financial Services

"Big Data" has no one definition. The National Institute of Standards and Technology (NIST) has defined big data in reference to the "Four Vs," as "consist[ing] of extensive datasets primarily in the characteristics of volume, velocity, variety, and/or variability that require a scalable architecture for efficient storage, manipulation, and analysis."[2] Each of these factors is important in how large data sets can be used effectively:

> (1) volume – the data sets are large and extensive;
> (2) velocity – data is generated, collected, and processed at a high rate, often in real time or near real time;
> (3) variety – different types of information can be used together in novel ways to draw inferences;
> (4) variability – this refers to changes in a data set, whether in the data flow rate, format/structure, or volume, that impacts its processing.

There has been widespread agreement that the use of big data "can produce tremendous benefits for society," as the FTC noted in its 2016 report on big data.[3] Large, sophisticated data sets can be used for a wide range of purposes in financial services. These range from purposes like fraud detection and compliance with anti-money laundering laws, to enabling better credit decisionmaking and providing consumers with financial management advice.[4] At an FTC hearing last year on algorithms, artificial intelligence, and predictive analytics, panelists discussed, for example, use of big data analytics to arrive at "fraud scores" that can help predict whether a transaction request is from someone other than the card holder,[5] as well as

---

[2] NATIONAL INSTITUTE FOR STANDARDS AND TECHNOLOGY, SPECIAL PUBLICATION 1500-1R2, NIST BIG DATA INTEROPERABILITY FRAMEWORK: VOLUME 1, DEFINITIONS 6, 11 (October 2019), *available at* https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-1r2.pdf.

[3] FEDERAL TRADE COMMISSION, BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION (January 2016) at 2, *available at* https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf ("FTC Big Data Report").

[4] For example, the FTC noted in its Big Data Report that "[c]ompanies have used big data to provide alternative ways to score populations that were previously deemed unscorable," and that "big data algorithms could help reveal underlying disparities in traditional credit markets and help companies serve creditworthy consumers from any background." FTC Big Data Report at 6-7.

[5] FEDERAL TRADE COMMISSION, TRANSCRIPT OF SEVENTH HEARING ON THE COMPETITION AND CONSUMER PROTECTION ISSUES OF ALGORITHMS, ARTIFICIAL INTELLIGENCE, AND PREDICTIVE ANALYSIS, PRESENTATION OF MELISSA MCSHERRY, SVP, GLOBAL HEAD OF DATA PRODUCTS, VISA, (Nov. 13, 2018) , *available at* https://www.ftc.gov/system/files/documents/public_events/1418693/ftc_hearings_session_7_transcript_day_1_11-13-18_0.pdf.

incorporation of new data sources into credit scoring.[6] Moreover, while sometimes called "alternative" or "non-traditional" data, many of these data sets can consist of information that already exists but has not been available to be used at scale for certain purposes (such as the example of cash flow data in lending decisions, as discussed further below).

The use of advanced data for credit decisionmaking is particularly promising. Large data sets can enable lenders to better analyze credit risk, and potentially expand access to credit to those who find it difficult to obtain credit when evaluated using traditional credit models. Many consumers are "thin file" or "no file" consumers who lack an adequate credit history to generate a reliable credit score, and others have relatively low scores that do not accurately reflect their current level of creditworthiness.[7]

The non-profit FinRegLab recently released the results of a promising study that illustrates the ability of large-scale data analytics to responsibly expand access to credit without raising issues related to bias. FinRegLab analyzed data from six non-bank financial service providers that used cash-flow information as part of their credit decisionmaking. Cash flow data can be obtained from consumer or small business accounts, and has the advantages of the "four Vs" of big data: substantial volume and variety of transactions, updated constantly.

The organization's study concluded that the "predictiveness of the cash-flow scores and attributes was generally at least as strong as the traditional credit scores and credit bureau attributes studied."[8] It also found that "participants appear to be serving substantial numbers of borrowers who may have historically faced constraints on their ability to access credit," and in regard to fair lending, that "the degree to which the cash-flow data predicted credit risk appeared to be relatively consistent across subpopulations" of race, ethnicity, and gender, and "appeared to provide independent predictive value across all groups rather than acting as proxies for demographic group."[9] FinRegLab also found the use of cash-flow data for credit underwriting appears to be spreading more rapidly in small business lending than in consumer lending, and that it is being used not only by online lenders, but also banks, payment processors, e-commerce platforms, and accounting service providers to provide small business loans.[10]

---

[6] FEDERAL TRADE COMMISSION, TRANSCRIPT OF SEVENTH HEARING ON THE COMPETITION AND CONSUMER PROTECTION ISSUES OF ALGORITHMS, ARTIFICIAL INTELLIGENCE, AND PREDICTIVE ANALYSIS, PRESENTATION OF ANGELA GRANGER, VP ANALYTICS, EXPERIAN 83 (Nov. 13, 2018), *available at* https://www.ftc.gov/system/files/documents/public_events/1418693/ftc_hearings_session_7_transcript_day_1_11-13-18_0.pdf.

[7] The CFPB has estimated that "26 million Americans are credit invisible, meaning they have no credit history with a nationwide consumer reporting agency [and] [a]nother estimated 19 million consumers have a credit history that has gone stale, or is insufficient to produce a credit score under most scoring models." Patrice Ficklin and Paul Watkins, *An update on credit access and the Bureau's first No-Action Letter*, CONSUMER FINANCIAL PROTECTION BUREAU BLOG (Aug. 6, 2019), https://www.consumerfinance.gov/about-us/blog/update-credit-access-and-no-action-letter/.

[8] *Fact Sheet: Cash-Flow Data In Credit Underwriting*, FINREGLAB, https://finreglab.org/fact-sheet-cash-flow-data-in-credit-underwriting/ (last visited Nov. 18, 2019).

[9] *Id.*

[10] *Id.*

Top officials at the Consumer Financial Protection Bureau (CFPB) also recently announced the results of the Bureau's data analysis conducted in connection with its no-action letter to Upstart Network, Inc. ("Upstart"). Upstart's underwriting model uses traditional underwriting data and various categories of alternative data, including information related to borrowers' education and employment history, and also uses machine learning in making credit underwriting and pricing decisions. The CFPB findings illustrate the benefits of using large data sets and machine learning to responsibly expand access to credit. It particular, the agency found:

- The company's tested model approved 27% more applicants than the traditional model, and yielded 16% lower average APRs for approved loans.

- This expansion of credit access reflected in the results occurred across all tested race, ethnicity, and sex segments, resulting in the tested model increasing acceptance rates by 23-29% and decreasing average APRs by 15-17%.

- In many consumer segments, the results showed that the tested model significantly expanded access to credit compared to the traditional model.

- As for fair lending concerns, when comparing the tested model with the traditional model, the approval rate and APR analysis results provided for minority, female, and 62 and older applicants showed no disparities that the CFPB found to require further fair lending analysis under the company's compliance plan.[11]

Regulatory landscape

Companies seeking to use large consumer data sets in financial services are currently subject to extensive regulation that governs how they deal with consumer data. Applicable federal laws include the Fair Credit Reporting Act (FCRA), Equal Credit Opportunity Act (ECOA), Gramm-Leach-Bliley Act (GLBA), and FTC guidance around data privacy and security. While the application of these laws may raise novel questions in some circumstances involving big data, and there may be opportunities to update or modernize them, financial services companies are already building in compliance as the volume and complexity of consumer data scale up.

FCRA. The FCRA, among other things, imposes obligations on consumer reporting agencies ("CRAs") that compile and sell defined "consumer reports" for purposes that include credit determinations. The FCRA requires CRAs to implement reasonable procedures to ensure "maximum possible accuracy" of consumer reports.[12] If a consumer files a dispute with a CRA, it must conduct a "reasonable investigation" as to the accuracy of the investigation.[13] And when creditors make certain adverse decisions based on consumer report information provided by a CRA, the creditor must provide notice to the consumer and information about the CRA that provided the consumer report.[14] In the context of big data, the FTC has said that "if an unaffiliated firm regularly evaluates companies' own data and provides the evaluations to the

[11] Ficklin and Watkins, *supra* note 7.
[12] 15 U.S.C. § 1681e.
[13] *Id*. § 1681i.
[14] *Id*. § 1681m.

companies for eligibility determinations, the unaffiliated firm would likely be acting as a CRA, each company would likely be a user of consumer reports, and all of these entities would be subject to Commission enforcement under the FCRA."[15]

ECOA. The ECOA prohibits discrimination on the basis of race, color, religion, national origin, sex, marital status, age, or receipt of public assistance.[16] In its Big Data Report, the FTC noted that ECOA applies to the use of big data analytics as well. So, for example, the report indicates that, in the FTC's view, "if a company makes credit decisions based on zip codes, it may be violating ECOA if the decisions have a disparate impact on a protected class and are not justified by a legitimate business necessity."[17] Additionally, when a creditor takes an adverse action, it must provide a consumer notification that includes an explanation of the reason for a decision.[18] The CFPB is also currently considering implementation of Section 1071 of the Dodd-Frank Act, which requires financial institutions to compile, maintain, and submit to the Bureau certain information concerning credit applications by women-owned, minority-owned, and small businesses.[19]

GLBA. The GLBA and its implementing regulations govern financial institutions' treatment of consumer data in connection with certain products or services.[20] The CFPB's implementing Regulation P, for example, requires covered financial institutions to provide certain privacy notices and to comply with certain limitations on the disclosure of nonpublic personal information to nonaffiliated third parties, and it requires financial institutions and others to comply with certain limitations on redisclosure and reuse.[21] The FTC's Safeguards Rule requires covered financial institutions to develop, implement, and maintain a comprehensive information security program, and the FTC is currently considering whether to amend the Rule to include more specific data security requirements.[22]

FTC privacy and data security actions. The FTC brings enforcement actions to protect consumer privacy and data security under Section 5 of the FTC Act, and its jurisdiction extends to non-bank financial technology companies.[23] The agency has published industry guidance based on its enforcement actions. In the area of data security, for example, it has outlined expectations in its *Start with Security* publication and *Stick with Security* blog post series.[24]

---

[15] FTC Big Data Report at 15.
[16] 15 U.S.C. § 1691 *et seq*.
[17] The report further notes that, "[e]ven if evidence shows the decisions are justified by a business necessity, if there is a less discriminatory alternative, the decisions may still violate ECOA." *Id*. at 19.
[18] 12 C.F.R. § 1002.9.
[19] *See CFPB Symposium: Section 1071 of the Dodd-Frank Act*, CFPB ARCHIVE OF PAST EVENTS (Nov. 13, 2019), https://www.consumerfinance.gov/about-us/events/archive-past-events/cfpb-symposium-section-1071-dodd-frank-act/.
[20] 15 U.S.C. § 6801 *et seq*.
[21] *See* 12 C.F.R. Part 1016.
[22] See 16 C.F.R. Part 314; Press Release, Federal Trade Commission, FTC Seeks Comment on Proposed Amendments to Safeguards and Privacy Rules (March 5, 2019), https://www.ftc.gov/news-events/press-releases/2019/03/ftc-seeks-comment-proposed-amendments-safeguards-privacy-rules.
[23] 15 U.S.C. § 45.
[24] *See* FEDERAL TRADE COMMISSION, START WITH SECURITY: A GUIDE FOR BUSINESS (June 2015) https://www.ftc.gov/system/files/documents/plain-language/pdf0205-startwithsecurity.pdf; *see also Stick with*

New privacy laws

In addition to the existing laws, discussed above, that regulate how financial services companies can use large consumer data sets, new cross-sectoral privacy laws have been enacted. These include the EU's GDPR, as well as state laws like the CCPA. These laws have imposed additional compliance costs on financial services companies and have resulted in regulatory uncertainty around how to handle particular types of consumer data.

GDPR. The GDPR regulates the collection and processing of personal data for individuals located in the European Union, among other things. The GDPR is based on six core principles, including being lawful, fair, and transparent and limiting the storage of data.[25] Under these six principles, there are a number of rights that data controllers (and to a lesser degree, data processors) must honor, and well as other business obligations. The GDPR provides individuals with a number of rights, including rights of access, rectification, deletion, and data portability, and the right to restrict data processing.[26] The GDPR also requires purpose specifications for collecting and processing personal data, which functions as a potential limitation on the use of large data sets that might be utilized for purposes beyond which the data was originally obtained.[27] The GDPR additionally includes a "data minimization" principle that limits processing of personal data to what is "adequate," relevant," and "necessary" in relation to the purposes for which it is processed.[28] Even companies with relatively small European operations or customers have incurred significant costs in coming into compliance with GDPR.

CCPA. The California Consumer Privacy Act is the most significant privacy law enacted by a state so far. It goes into effect on January 1, 2020. The law applies to businesses that collect and control the processing of California residents' personal information, do business in the state, and meet certain qualifications in terms of revenues or data collection.[29] The "personal information" covered broadly includes traditional identifiers, commercial information, biometric information, unique personal identifiers (like IP addresses or cookies), internet information like browsing history or geolocation data, and inferences drawn from any information to create a profile about a consumer.[30]

The law has created a number of compliance challenges for businesses. First, the substantive requirements of the law have been a moving target, and significant uncertainty remains about how to operationalize a complex and often unclear law, even though it will become effective in less than two months. Amendments to the law were passed and signed into law as late as October 11, 2019. On October 10, 2019, the California Attorney General released extensive

---

Security: A Business Blog Series, FTC BLOG (October 2017), https://www.ftc.gov/tips-advice/business-center/guidance/stick-security-business-blog-series.

[25] *See* Eur. Par. And Council Regulation 2016/679 of Apr. 27, 2016 Protection of Natural Persons With Regard to the Processing of Personal Data and the Free Movement of Such Data, and repealing Direction 95/46/EC, art. 5 ("GDPR").

[26] *See* GDPR art. 15-18, 20.

[27] *See* GDPR art. 5(1)(b).

[28] *See* GDPR art. 5(1)(c).

[29] *See* California Consumer Privacy Act of 2018 ("CCPA"), Cal. Civ. Code § 1798.140(c) (2018).

[30] *Id.* § 1798.140(o).

draft regulations to implement the law, and many of these go beyond what is required in the law itself and are themselves ambiguous. Adding to the lack of clarity, we are in the middle of a two-month comment period regarding the draft regulations, which closes on December 6. The final regulations must be adopted by July 1, 2020. All of this creates a period of uncertainty and raises practical burdens for companies attempting to comply with the law. While enforcement of the law is delayed until July 1, 2020 or six months after the Attorney General adopts implementing regulations, companies are striving to put procedures in place for compliance by January 1 when the law goes into effect, but the specifics of many of the procedures governed by the draft regulations remain subject to change. Additionally, starting on January 1, covered consumers will be able to seek information about the collection of their personal information, and the disclosure or sale of their information to third parties that occurred over the past year.[31] That means that, even now, before the law has gone into effect, we are in "look back" period where companies will have obligations for providing information about their use of personal information.

Second, the CCPA creates a patchwork of rules potentially applicable to financial institutions. Section 1798.145(e) states that the CCPA "shall not apply to personal information collected, processed, sold, or disclosed pursuant to the federal Gramm-Leach-Bliley Act [GLBA] . . . . and implementing regulations."[32] As noted above, however, the GLBA applies to defined "financial institutions" and covers only certain personal information that is provided by an individual, results from a transaction with or service performed for an individual, or is otherwise obtained by a financial institution, in connection with a financial product or service to be used primarily for personal, family, or household purposes.[33] As a result, some of the personal information that financial services providers collect, or collect for certain purposes, arguably may not be covered by GLBA and its implementing regulations and may covered by the CCPA. Financial services providers may consider implementing new procedures – including notices and procedures for responding to and verifying consumer requests – that may apply only to certain data.

Third, the law imposes significant compliance burdens on companies that do business nationwide – while leaving open the possibility that other states could pass laws that go even further or that may be inconsistent with California's mandates. This patchwork approach is confusing for both consumers and companies trying to comply with the law. These burdens will be felt by businesses of all sizes, and may be particularly problematic for small businesses, which may be covered because they deal with substantial amounts of consumer data.[34] Moreover, companies must build in compliance with the current law with an eye on what other state legislatures may pass in the near future.

Additionally, some observers have suggested that the CCPA may inhibit the ability to effectively collect and use large data sets for purposes of implementing machine learning models. A recent report by the American Bar Association Section of Antitrust notes that, if a significant number of

[31] *Id.* § 1798.100(d).
[32] *Id.* § 1798.145(e).
[33] *See* 15 U.S.C. § 6809(3), (4)(A), (9).
[34] Cal. Civ. Code § 1798.140(c)(1)(B).

consumers were to exercise their deletion rights in certain circumstances, that might result in data sets that are not representative of the relevant population.[35] The ABA report also raises the possibility that "it may be difficult for a company to specify at or before the point of collection the purposes for which the business will use the data in the context of analytics."[36] Whether this proves true in practice, companies will want to think carefully about how they define the purposes for which they collect consumer data that may be incorporated into larger data sets to enable beneficial services to be provided to consumers.

++++++

Financial services companies are currently making significant advances in expanding financial services for the benefit of consumers. In evaluating current or proposed privacy laws in the context of "big data," it is important to weigh any purported benefits against the significant benefits from innovation in financial services that consumers also want.[37] Particularly in financial services, we should recognize that new regulations have the potential to burden important pro-consumer innovation that can materially improve consumers' financial lives.

Thank you again, and I look forward to answering your questions.

[35] AMERICAN BAR ASSOCIATION SECTION OF ANTITRUST, ARTIFICIAL INTELLIGENCE & MACHINE LEARNING: EMERGING LEGAL AND SELF-REGULATORY CONSIDERATIONS 59 (Sep. 30, 2019), *available at* https://www.americanbar.org/content/dam/aba/administrative/antitrust_law/comments/october-2019/clean-antitrust-ai-report-pt1-093019.pdf.

[36] *Id*. at 57.

[37] In the broader context of U.S. privacy law, the U.S. Chamber of Commerce has released privacy principles that can be found at https://www.uschamber.com/sites/default/files/9.6.18_us_chamber_-_ctec_privacy_principles.pdf.