July 1, 2021

Chief Counsel's Office,
Attention: Comment Processing,
Office of the Comptroller of the Currency,
400 7th Street SW, Suite 3E-218,
Washington, DC 20219

James P. Sheesley, Assistant Executive Secretary,
Attention: Comments-RIN 3064-ZA24,
Federal Deposit Insurance Corporation,
550 17th Street NW, Washington, DC 20429

Ann E. Misback, Secretary,
Board of Governors of the Federal Reserve System,
20th Street and Constitution Avenue NW,
Washington, DC 20551

Bureau of Consumer Financial Protection,
1700 G Street NW,
Washington, DC 20552.

Melane Conyers-Ausbrooks, Secretary of the Board,
National Credit Union Administration,
1775 Duke Street, Alexandria,
VA. 22314-3428

Re: **Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning (Docket Nos: OCC-2020-009, OP-1743, CFPB-2021-0004, NCUA-2021-0023, RIN 3064-ZA2)**

Dear Madam or Sir:

Upstart Network, Inc. ("Upstart") provides technology services to financial institutions to enable them to lend to consumers online. Upstart's credit underwriting platform, now eight years old, harnesses artificial intelligence ("AI") and machine learning ("ML") and uses data that goes beyond traditional credit scores, helping financial institutions of all sizes identify creditworthy consumers online and price risk more accurately. Well regulated partnerships between financial institutions and technology companies like Upstart are critically important today for the financial health of consumers and the banking system. More and more consumers are seeking credit and applying for loans from their mobile

devices or home computers, while physical branch networks continue to shrink, often placing a burden on low and moderate income areas, rural areas and communities of color.[1]

Working with Upstart can help banks do more than convert a traditional loan product into a digital offering. Because of Upstart's use of additional data and AI/ML techniques, the banks and credit unions that work with Upstart are able to offer loans to more consumers who might not qualify using traditional underwriting methods. Among other things, this enables them to increase the percentage of consumer loans that are made to low-and-moderate income borrowers.[2] It is also critically important that the use of AI leads to fair outcomes, and Upstart has worked proactively with the CFPB to demonstrate that using AI technology in lending can improve credit access and reduce interest rates for borrowers in all demographic groups, when compared to traditional underwriting approaches.[3]

This request for information on AI and financial institutions is an important step forward. Access to fair, affordable credit is critical for economic mobility.[4] Regulators of financial institutions must help ensure that America's consumers can realize the benefits from AI lending in a well regulated, and supervised context. Upstart believes the agencies' joint request for information ("RFI") is a promising step toward achieving that goal. Our response includes specific suggestions for agency action. In particular, we urge that the agencies, through coordinated action, work to revise, clarify, and update their existing model risk management, fair lending, and third-party oversight guidance so that banks and their service providers can understand and comply with expectations for the use of AI/ML in consumer lending.

Question 1: How do financial institutions identify and manage risks relating to AI explainability? What barriers or challenges for explainability exist for developing, adopting, and managing AI?

As Upstart has worked to develop its AI credit underwriting model, variables are selected and used in the model on the basis of their ability to more accurately predict default. The volume and variety of the data sets used by Upstart's models expand the number of applicants bank partners can approve and

---

[1] National Community Reinvestment Coalition Bank Branch Closure Update 2017-2020. https://ncrc.org/research-brief-bank-branch-closure-update-2017-2020/

[2] Through March 31, 2020 45.5% of loans made relying on the Upstart model go to individuals who would meet the definition of being low or moderate income. LMI calculations in this internal analysis are approximate using Upstart borrower data: reported individual borrowers' income were used in lieu of household income and zip codes were used as a proxy for census tract information. Upstart By The Numbers. https://www.upstart.com/blog/upstart-by-the-numbers

[3] An update on credit access and the Bureau's first No-Action Letter. https://www.consumerfinance.gov/about-us/blog/update-credit-access-and-no-action-letter/

[4] https://equitablegrowth.org/race-and-the-lack-of-intergenerational-economic-mobility-in-the-united-states/

promote more inclusive lending; the complexity involved in identifying predictive, nonobvious associations necessitates Upstart's use of AI/ML algorithms. The accuracy gains from these techniques allow the Upstart model to discover far more consumers who would not be considered creditworthy by the traditional credit scoring system, and the ability to price those consumers more accurately offers the opportunity to provide them with more favorable terms. Indeed, four out of five Americans who have taken out a loan have never defaulted, yet less than half of Americans have access to prime credit.[5]

While AI offers a significant opportunity to improve the accuracy, fairness, and inclusiveness of the models used by financial institutions, Upstart believes it is also critical that AI model outputs meet a basic standard of being "explainable." The largest challenge for explainability of AI systems is the fact that they often make a large number of decisions before reaching their final output. Today there is a growing body of both established tools, and newer promising approaches, that can facilitate interpretation of complex AI and machine-learning models.[6] These include, for instance:

1. Shapley values (SHAP)[7]
2. Partial-dependence plots[8]
3. Relative importance[9]
4. Permuted feature importance[10]
5. Individual conditional expectation[11]
6. Local interpretable model-agnostic explanations (LIME)[12]

These techniques, either used separately or in combination, offer financial institutions ways to quantify the impact of particular data sets and even individual variables in a model. For example, one of the techniques Upstart uses is SHAP, which enables it to quantify the impact of certain variables on

---

[5] According to an Upstart retrospective study completed in December 2019. This study defined access to prime credit as individuals with credit reports with VantageScores of 720 or above.

[6] Papastefanopoulos, V., Kotsiantis, S. (2020). "Explainable AI: A Review of Machine Learning Interpretability Methods". Entropy, (2021), 23, 18. https://dx.doi.org/10.3390/e23010018

[7] Lundberg, S., Lee, S.I. "A Unified Approach to Interpreting Model Predictions". In: NIPS (2017). papers.nips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

[8] Friedman, J. (2001). "Greedy Function Approximation: A Gradient Boosting Machine". The Annals of Statistics, 29:5, 1189-1232. *See also*, Goldstein, A., Kapelner, A., Bleich, J., Pitkin, E. (2015). "Peeking Inside the Black Box: Visualizing Statistical Learning with Plots of Individual Conditional Expectation". Journal of Computational and Graphical Statistics, 24:1, 44-65. https://doi.org/10.1080/10618600.2014.907095

[9] Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1984). C. (1984). Classification and Regression Trees, Wadsworth, New York.

[10] Fisher, A., Rudin, C., Dominici, F. (2019). "All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously". Journal of Machine Learning Research, 20:177, 1-81. jmlr.org/papers/volume20/18-760-18-760.pdf

[11] Goldstein, A., Kapelner, A., Bleich J., Pitkin, E. (2015), "Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation". Journal of Computational and Graphical Statistics, 24:1, 44-65, DOI: 10.1080/10618600.2014.907095

[12] Ribeiro, M.T., Singh, S., Guestrin, C. (2016). "Why Should I Trust You?: Explaining the Predictions of Any Classifier". KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

model outputs. The SHAP approach quantifies this impact by assessing the effect of a variable's removal from the model. This method allows each variable to be assigned some fraction of the model's output for which it is accountable (variables with a negative impact are assigned a negative proportion of this effect). In the context of Upstart's underwriting model, this method allows Upstart to ascertain which variables or variable groupings are most influential in producing credit decisions for a particular applicant. This method has numerous business applications including the determination of reasons that are included in adverse action notices ("AANs"). By observing the model's reliance not only on discrete variables but also of highly correlated variable groupings, more specific and nuanced adverse action reasons can be communicated to a consumer.

Upstart has found that to manage the risks associated with AI, financial institutions require rigorous testing and third party validation of AI model outputs in combination with the use of available tools and techniques to ensure accurate and relevant explanations of those final outputs. For example, during model development at Upstart, data scientists follow a rigorous statistical process of cross-validation to ensure that every added variable produces a robust improvement in model accuracy and the causal relationships are well understood. All testing is fully documented before any updated model is ready to be put into production.

Despite the growing body of explainability techniques outlined above, misplaced perceptions about insufficient "explainability" tools may delay the deployment of AI / ML systems in the banking system.[13] Institutional inertia or status quo bias may cause sluggish adoption of AI/ML applications, even if traditional models are less accurate and exhibit problems with interpretability.[14] These forces could impede the deployment of sound AI credit underwriting systems at financial institutions of all sizes, making these institutions less competitive.

Furthermore, in the case of credit underwriting models, these perceptions could become barriers to the availability of affordable consumer credit. Access to insured deposit funding typically means that banks are "...the most dependable, low cost, through-the-cycle source of credit for consumers, including LMI borrowers."[15] Fewer bank AI-powered choices will mean higher fees, higher interest rates, or, in certain cases, a lack of affordable access.

---

[13] For example, the Financial Stability Board released a report in 2017 assessing the costs and benefits of AI use in financial services. Lack of interpretability was identified as a key risk, with the claim that it could result in unpredictable and unforeseen actions with possible macroeconomic consequences. See: Financial Stability Board (FSB), Artificial intelligence and machine learning in financial services. Market developments and financial stability implications. Nov. 1, 2017, available at: https://www.fsb.org/2017/11/artificial-intelligence-and-machine-learning-in-financial-service/.

[14] Traditional credit models can suffer from problems with "interpretability." For example, an input variable in a simple multivariate regression model could be found empirically to influence default risk in a manner that is difficult to understand or rationalize.

[15] Bank Policy Institute and Covington, Artificial Intelligence Discussion Draft: The Future of Credit Underwriting: Artificial Intelligence and Its Role in Consumer Credit (2019) at p. 6.

Question 2: How do financial institutions use post-hoc methods to assist in evaluating conceptual soundness? How common are these methods? Are there limitations of these methods (whether to explain an AI approach's overall operation or to explain a specific prediction or categorization)? If so, please provide details on such limitations.

AI model developers should provide financial institutions with the tools required to do rigorous ongoing post-hoc evaluation and oversight of the model and its performance. Upstart provides these types of tools to financial institutions, including access to proprietary dashboards and a reporting API that delivers loan level data directly to the lender in real time allowing the lender to track performance of its lending program and ingest data for analytical purposes.[16]

Post-hoc methods are critical for evaluating soundness of models. Monitoring the performance of models after deployment provides the best "out-of-sample" test of model accuracy and performance. For example, Upstart monitors the performance of its underwriting model by comparing observed outcomes with predicted loan outcomes (e.g., default rates). Periodic third-party validations of AI models for accuracy and fairness are also critical in providing additional, independent assessments of model performance to supplement the financial institution's internal monitoring process.

Notwithstanding these benefits, one limitation of post-hoc methods is that it may be difficult initially to pinpoint the cause of model underperformance. For example, model underperformance in the context of consumer loans may arise as a result of a changing borrower behavior, macroeconomic conditions, or competition from other financial institutions. As such, it may be difficult to adjust models quickly to correct for underperformance. However, it is critical to note that these problems also exist with traditional static credit models.

Frequent updating of AI models can also offer an advantage over traditional static lending models because they can adapt more readily to changing economic conditions. A potential limitation also occurs when AI models are updated frequently, if each model update is only tested on a limited sample of observations before the model is updated again. This limitation can be fully addressed, however, by also employing validation methods that use past data (e.g., cross-validation) to ensure that these factors do not hinder model performance.

Question 3: For which uses of AI is lack of explainabilitymore of a challenge? Please describe those challenges in detail. How do financial institutions account for and manage the varied challenges and risks posed by different uses?

[16] Banks that use Upstart's platform are provided with the various dashboards and a reporting API to track their raw lending data and other outputs in real time.

As noted in Question 1, general discussions surrounding any perceived "explainability" or "interpretability" shortcomings of AI[17] often do not consider the various actual precise use cases - i.e. the types of explanations that are required of financial institutions, or the existing and emerging tools and techniques that can help to deliver those explanations (discussed in Question 1).[18] Unfortunately, this could result in a slower pace of adoption of AI models by financial institutions, with implications for both lending accuracy and financial inclusion.

As noted frequently in this RFI response, there are many different potential applications of AI in financial institutions, with different explainability requirements. If an AI-system is being employed to detect fraudulent loan applications for instance, it may be sufficient for the financial institution to continually review the system for accuracy and to have basic explanations for why the anti-fraud system is, or is not, performing adequately in its role. We discuss in detail fair lending explainability requirements and effective model oversight in our responses to Questions 9, 10 and 15.

In Upstart's case, bank and credit union customers use Upstart largely for its AI credit decision making model, which harnesses the power of AI and machine learning to help these financial institutions determine creditworthiness and fair pricing of applicants. One particular explainability challenge in AI lending includes the fact that generating accurate explanations from AI lending model outputs may be time consuming, requiring a variety of comprehensive testing, evaluation, and reporting for both model accuracy and differing dimensions or definitions of fairness. Upstart has therefore put significant effort into ensuring that the AI models it has put into production for financial institutions are explainable to (1) consumers who apply for loans, (2) the financial institutions that leverage the AI model, and (3) the regulators that oversee financial institutions and activities for safety and soundness and for consumer protection.

To manage these issues effectively, Upstart and its financial institution partners tailor the tools and approaches used based on the respective audience. For instance, a loan borrower is interested in a simple explanation of the relevant reason(s) that they received a negative outcome from an AI model, and whether there is anything they can do to change their circumstances to receive a better outcome in the future. Here, Upstart uses techniques outlined in Question 1. Model developers, regulators and compliance teams will be interested in those issues, but they will also be interested in the more granular operation of the AI system and a deeper level of technical evidence to confirm the model remains statistically sound and fair.[19] Further, prudential regulators may also require clear evidence that the model is performing accurately and fairly (and/or explanations why it is not), while the CFPB's may focus on fair lending and ensuring that AI/ML models used in credit underwriting and customer

[17] Supra note 13
[18] For a representative list, see response to question 1.
[19] This process may include a model validation audit by a recognized and certified third party.

relationship management do not produce biased or discriminatory outcomes and are being tested regularly and strenuously to detect for such outcomes.

In conclusion, Upstart believes that the AI model outputs that financial institutions use to make lending decisions should be explainable.[20] To manage risks related to AI explainability from the outset, AI systems should only be put into production when human operators have sufficient confidence in the system and know that the system is reliably operating and that they can vigilantly monitor and explain its decisions to all of the different audiences.

Question 4: How do financial institutions using AI manage risks related to data quality and data processing? How, if at all, have control processesor automated data quality routines changed to address the data quality needs of AI? Howdoes risk management for alternative data compare to that of traditional data? Are thereany barriers or challenges that data quality and data processing pose for developing, adopting, and managing AI? If so, please provide details on those barriers or challenges.

In the joint regulators' December 2019 *"Interagency Statement on the Use of Alternative Data in Credit Underwriting"* federal financial regulators laid out the opportunities and considerations for financial institutions looking to manage alternative data use.[21] Effective AI models rely on very large amounts of data. For example, Upstart's underwriting model incorporates more than 1,000 variables (e.g., credit history variables, loan duration and amount, income and employment variables, etc.) and benefits from a rapidly growing training dataset that contains more than 10.5 million repayment events.[22]

When incorporating sources of data into AI / ML models, Upstart performs extensive testing and validation when evaluating data quality from all data sources. Upstart performs automated testing and employs a rigorous change management procedure before making any changes to the production code or modeling. To develop its model, Upstart has leveraged mature processes and best practices from the software engineering industry and applied them to AI applications. For example, Upstart extensively tests AI models to ensure their outputs are what were predicted, including employing post-hoc dashboards with in-depth reporting, and regular monitoring of model behavior and performance.

Upstart's management of risks related to data quality and processing are similar for alternative and traditional credit data it uses. For example, program code may produce automated error messages when data inputs have apparent irregularities in terms of format or magnitude. However, there may be more custom work required with alternative data, which generally also involves more specialized testing. There are specific data challenges with respect to data entered by applicants, in particular, as

[20] See response to question 1
[21] *See* https://www.ncua.gov/files/press-releases-news/alternative-data-use-credit-underwriting.pdf
[22] This data is as of March 31, 2021.

inaccurate information could be input either intentionally or unintentionally. Upstart has developed a verification model that flags applications for risk of fraud or misinformation, which can then be either automatically verified through consumer reporting agencies, employers, universities, bank accounts, and other third parties and third party databases or manually verified through additional documentation provided by the applicant.[23]

To fully mitigate these risks, Upstart engages in frequent model testing, using both human reviews and automated testing tools, to ensure that all parts of the model are performing as intended, with testing tailored to the underlying data / use involved.

Question 5: Are there specific uses of AI for whichalternative data are particularly effective?

Upstart has demonstrated that helping financial institutions lend money safely to more consumers, or alternatively, reduce their credit losses, is an area uniquely suited to the use of AI and alternative data.[24] In 2017, the CFPB "estimate[d] that 26 million Americans are 'credit invisible,' meaning they have no credit history at all," and "another 19 million people have credit histories that are too limited or have been inactive for too long to generate a credit score" under traditional credit scoring models.[25] Furthermore, CFPB has found that Black and Hispanic Americans are more likely than white or Asian Americans to be credit invisible or to have un-scored records and typical approaches to building strong credit files -- for example, "[t]he use of co-borrowers and authorized user account status -- [are] notably less common in lower-income neighborhoods."[26] To promote fair access to credit for all individuals, including those in these circumstances, federal regulators have recognized that credit underwriting is an area where AI/ML and their use of alternative data can be particularly effective. The CFPB, for instance, has reported that:

> *"In addition to the use of alternative data, increased computing power and the expanded use of machine learning can potentially identify relationships not otherwise discoverable through methods that have been traditionally used in credit scoring. As a result of these innovations, some consumers who now cannot obtain favorably priced credit may see increased credit access or lower borrowing costs."*[27]

---

[23] Given the magnitude of data used by AI models versus traditional underwriting techniques, it is possible that undetected data errors or corruption during storage or transmission may be harder to find, and more rigorous and diligent controls may be required.

[24] See Upstart.com Results To Date. https://www.upstart.com/about#results-to-date-3

[25] Schmidt & Stephens, supra note 8, at 141-142. 24 See id. 25 Cordray, supra note 1; see also Kenneth P. Brevoort, et al., CFPB Office of Research, "Data Point: Credit Invisibles," and https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf

[26] See Kenneth P. Brevoort & Michelle Kambara, CFPB Office of Research, "Data Point: Becoming Credit Visible," *available at* https://files.consumerfinance.gov/f/documents/BecomingCreditVisible_Data_Point_Final.pdf.

[27] *"An Update on credit access and the Bureau's first No-Action Letter,"* CFPB Blog (Aug. 6, 2019), *available at* https://www.consumerfinance.gov/about-us/blog/update-credit-access-and-no-action-letter/.

Upstart's AI/ML-powered credit underwriting model has demonstrated how AI/ML technology and alternative data use can significantly expand fair credit availability in a fair and responsible manner, consistent with fair lending laws and regulations, while also potentially improving bank safety and soundness.[28] The model has been proven to (1) be more accurate than traditional underwriting models, and (2) allow for greater access to credit at lower rates of loss.[29] Upstart's validation testing has consistently shown, using measures of statistical accuracy like AUC, that the model improves accuracy by at least two times compared to that of a traditional model.[30] An internal study comparing the Upstart model to that of several large U.S. banks found that our model could enable these banks to lower loss rates by almost 75% while keeping approval rates constant.[31] Further, an access-to-credit review by Upstart of its 2020 data using comparison methodology specified by the CFPB, showed that our AI model approved 26% more borrowers than high-quality traditional lending models at 10% lower APRs.[32] Further, the automation of the loan applications, including the underwriting process, using AI/ML technology provides a more streamlined and efficient process that benefits both financial institutions and consumers; approximately 70% of loans originated through Upstart's platform by bank partners are fully automated.

There also are fair lending benefits to AI/ML credit underwriting models compared to traditional credit underwriting models. Upstart's quarterly fair lending test results to date indicate that AI models can be used without generating unlawful disparate treatment of, or disparate impact on, protected-class borrowers. In addition to the quantifiable benefits of AI/ML credit underwriting models, the results also demonstrate the qualitative benefits to this technology as compared to traditional credit underwriting models. In particular, use of AI/ML technology can help eliminate unconscious or conscious human bias in the credit underwriting process through the use of AI/ML credit underwriting models that require little, if any, human intervention.

---

[28] Regulators have consistently identified the importance of more accurate credit underwriting for safety and soundness of financial institutions:
https://www.minneapolisfed.org/article/2014/underwriting-standards-lessons-from-the-past

[29] https://www.upstart.com/about#results-to-date-3

[30] Based on an internal studies comparing Upstart's model with a hypothetical lending model formulated using Upstart's approximation of credit score variables used in traditional simple rules-based lending models and additional variables including loan amount, debt-to-income ratio, monthly income, number of credit inquiries and number of trade accounts.

[31] In an internal study, Upstart replicated three bank models using their respective underwriting policies and evaluated their hypothetical loss rates and approval rates using Upstart's applicant base in late 2017. To compare the hypothetical loss rates between Upstart's model and each of the replicated bank models, Upstart held approval rates constant at the rate called for by each bank's respective underwriting policy. The results represent the average rate of improvement exhibited by Upstart's platform against each of the three respective bank models.

[32] Approval numbers compare the 2020 loan approval rate by the Upstart model and a hypothetical traditional credit decision model. The APR calculation compares the two models based on the average APR offered to borrowers up to the same approval rate. The hypothetical traditional model used in Upstart's analyses was developed in connection with the access-to-credit reporting requirements under its CFPB No- Action Letters, is trained on Upstart platform data, uses logistic regression and considers traditional application and credit file variables.

While these quantitative and qualitative benefits of AI/ML credit underwriting models make a compelling case for use of this technology in any economic environment, the COVID-19 pandemic has intensified the need for the critical role that AI/ML credit underwriting models can play during an economic recovery. For example, Upstart has experienced only half the increase in payment impairments – both at the peak of COVID-19 pandemic, and to date – compared to the industry standard, despite the fact that borrowers of Upstart-driven loans have a 25 point lower average FICO score.[33] A significant part of the success of Upstart's AI/ML credit underwriting model under these circumstances can be attributed to the model's ability to quickly incorporate unemployment data and other economic forecasts.

Question 6: How do financial institutions manage AI risks relating to overfitting? What barriers or challenges, if any, does overfitting pose for developing, adopting, and managing AI? How do financial institutions develop their AI so that it will adapt to new and potentially different populations (outside of the test and training data)?

Financial institutions can manage challenges related to "overfitting" via several methods. Overfitting errors happen when, for instance, AI/ML models identify, incorporate, and start to consistently make assumptions based on patterns that exist within a particular, narrow dataset that are not generalizable outside of that data set. Overfitting errors, left unaddressed, could lead to inaccurate or unfair model performance, or both.[34] Financial institutions should use AI technology from a vendor/supplier that has rigorous approaches that can help the institution avoid any overfitting.

There are a number of techniques that can be used to prevent overfitting in an AI model's operation. One effective technique is the use of cross-validation and hyperparameters. The cross-validation method divides datasets into two parts - a "training" dataset on which a model is estimated or trained and a "test" dataset on which the estimated model is evaluated. This method constitutes a form of out-of-sample testing. This kind of out-of-sample testing reduces overfitting because it ensures that the model fits well on data outside that on which it was originally estimated. Hyperparameters are used to help ensure the accuracy of the model when applied to the test dataset. These hyperparameters usually assign a penalty to models with high complexity or numbers of parameters to prevent overfitting to the training dataset.

---

[33] Based on a comparison of Upstart payment impairment rates to industry impairment rates provided in "dv01 Insights COVID-19 Performance Report Volume X," dated as of March 31, 2021 (the "dv01 Report"). The dv01 Report analyzed over 2.5million active loans with a weighted average FICO score of 720, which is 50 points higher than the weighted average FICO score of Upstart borrowers. Payment impairments include both hardships and delinquencies.

[34] A publicized example of overfitting occurred with the estimation of earthquake risk prior to the Fukushima nuclear disaster in 2011. Estimating the relationship between earthquake frequency and magnitude using two connected lines resulted in a possible underestimation of risk prior to the event because of statistical fluctuations or errors in the historical data. Using a simpler model consisting of a single line would have resulted in a higher estimate of risk prior to the event. See Nate Silver, "The Signal and the Noise: Why So Many Predictions Fail – but Some Don't. (2015).

In order to identify and prevent overfitting errors that lead to bias in models, financial institutions should conduct frequent fair lending tests to identify if any groups are underrepresented in testing or are treated unfairly by AI models, and address any disparities when they are identified (see responses to Questions 11-15). Last, improving the integrity and quality of the datasets used by a model must be an ongoing endeavor for all responsible AI model operators. Many smaller banks may not have internal teams that are able to do this work themselves; they should be able to feel comfortable relying on partners that have this capability.

Finally, Upstart and other financial institutions should continually seek to better understand and improve lending for borrowers at the margins of approval. This effort can increase the representation of underrepresented groups in their datasets and ensure their training data and testing adapts to changing populations.

Question 7: Have financial institutions identified particular cybersecurity risks or experienced such incidents with respect to AI? If so, what practices are financial institutions using to manage cybersecurity risks related to AI? Please describe any barriers or challenges to the use of AI associated with cybersecurity risks.Are there specific information security or cybersecurity controls that can be applied to AI?

Developing secure products and processes must be a top priority for businesses that handle large volumes of sensitive data, regardless of whether they use AI tools. Upstart develops information security protocols by design, with a robust development process framework built around the security principles of authentication, authorization, encryption, logging, and monitoring. AI / ML models typically enable operation on large data sets and therefore, cyber-security protocols should be appropriate for the size and type of the data sets in use and the processes they run. It's important to note that the use of AI models does not in and of itself create significantly increased cyber security risks for financial institutions because typically data that is used in AI models is completely depersonalized and as such, despite the sheer quantity of data used by an AI models, this does not correlate into significantly increased cybersecurity risk.

Federal and state law places a high bar for the handling of the data used by financial institutions to underwrite credit in any system, and those standards are applied rigorously to AI models.[35] Companies harnessing AI must use a strong information security infrastructure, detection tools, and oversight to

[35] 15 U.S.C. § 6801(b); *Interagency Guidelines Establishing Information Security Standards,* 66 Fed. Reg. 8616 (Feb. 1, 2001) and 69 Fed. Reg. 77610 (Dec. 28, 2004) promulgating and amending 12 C.F.R. Part 30, app. B (OCC); 12 C.F.R. Part 208, app. D-2 and Part 225, app. F (Board); 12 C.F.R. Part 364, app. B (FDIC); and 12 C.F.R. Part 570, app. B (OTS); Federal Trade Commission's Safeguards Rule, 16 C.F.R. part 364; Federal Financial Institutions Examination Council (FFIEC) Information Technology Examination Handbook's Information Security Booklet.

support the volume of sensitive data they rely on. All financial institutions, whether they are using AI or traditional models, are reviewed by regulators and/or other independent third parties regularly for data integrity, secure software principles, data accuracy, fairness, and risk. This includes ensuring that "confidentiality, integrity and availability" are preserved and also making sure that external actors have not penetrated the system and compromised an AI model's accurate functioning or in any way tainted the quality of the underlying data used in testing or production. It's important to note that third party vendors to financial institutions must operate in accordance with all applicable laws, including the Bank Service Company Act, which empowers regulators to examine the performance of services by certain third parties as though the services were being performed by the bank itself[36] -- including for sound cybersecurity protocols.

Question 8: How do financial institutions manage AI risks relating to dynamic updating? Describe any barriers or challenges that may impede the use of AI that involve dynamic updating. How do financial institutions gain an understanding of whether AI approaches producing different outputs over time based on the same inputs are operating as intended?

To manage AI risks related to dynamic updating, it is critical that meaningful monitoring and controls are implemented and regular testing is conducted to review model behavior. This mitigates the risk that any update causes the model to work in an unintended manner, that for instance, could lead to overfitting or other problematic outcomes. In our observation, models may change dynamically over time for at least two reasons: first, model parameters can be updated when new data are included in the training dataset and second, AI modelers can change the code and algorithms to improve model performance.

To ensure that its AI models do not work in unintended ways, Upstart has an active program of regularly evaluating new sources of data to use as inputs for its models. In addition, Upstart continually assesses whether the model's predictions fit observed data from its loan portfolio over time. Upstart does this assessment initially during model development when it uses cross-validation methods to determine whether a potential change improves model accuracy metrics. It also does this assessment on a periodic long-term basis in its evaluation of predicted versus observed defaults, rates of return, and losses, overseen by internal committees. Finally, Upstart engages in higher frequency monitoring of lending metrics such as loan sizes, approval rates, and conversion rates to identify any possible anomalies with model updates.

Question 9: Do community institutions face particular challenges in developing, adopting, and using AI? If so, please provide detail about suchchallenges. What practices are employed to address those impediments or challenges?

[36] 12 U.S.C. § 1867(c).

As an AI model developer that partners with community banks, Upstart has experienced first-hand situations where some community institutions feel they are unable to access the benefits of this technology due to challenges such as uncertainties over their own technical expertise, regulatory compliance obligations, and/or perceptions that they have inadequate expertise, or human and financial resources, to fulfill those obligations.[37] Still, small banks and credit unions represent the majority of the roughly 20 institutions that use Upstart's technology, indicating that these challenges and perceptions are not uniform impediments across the system.

Upstart has found that certain best practices are key to addressing the challenges faced by small and medium-sized community institutions in deploying AI, including: (1) offering to participate transparently in discussions with their regulatory supervisors; (2) conducting regular independent statistical validations of the Upstart model that institutions can review; (3) providing access to loan data and reports that align with the regulatory examination schedule of the institution; and (4) articulating clear strategic goals early in the onboarding process for how the deployment of a third-party AI credit model will help the institution execute on its business strategy and better serve the community.

Upstart notes that Section 7 of the Bank Service Company Act ("BSCA"), in addition to requiring depository institutions to notify their respective federal banking agency of contracts or relationships with service providers, also subjects the performance of such services by service providers to regulation and examination by the federal banking agencies *to the same extent as if the services were being performed by the depository institution.*[38] The BSCA framework requires service providers to comply with the institution's regulatory standards and ensures they are subject to the oversight of the institution's examiners. As such, federal regulators should take steps to ensure that community institutions and their examiners understand that these institutions may be permitted to rely on the expertise of its third party service providers (such as Upstart) for related compliance matters, including, for instance, when answering questions posed by regulators / supervisors related to their activities/services.[39]

Finally, development of a well-designed and well-executed standard-setting process and voluntary certification of third-party models may provide an opportunity for regulators to ease the path to adoption of sound third party models. This enables community banks to use these AI models in a safe and sound manner even when they do not have adequate resources to independently validate AI models developed by third-party technology providers themselves. Upstart believes that the recent

---

[37] Community institutions often express misgivings about whether they will be able to answer very technical model documentation questions from supervisors / regulators and express uncertainty as to whether they can rely on third party expertise to assist them with monitoring and oversight.

[38] 12 U.S.C. § 1867(c). *See also* https://www.aba.com/banking-topics/compliance/acts/bank-service-company-act

[39] Federal law has long encouraged technical assistance for certain small institutions. One example is Section 308 of the Financial Institutions Reform, Recovery, and Enforcement Act of 1989 (FIRREA) which established several goals related to minority depository institutions (MDIs), including "providing for training, technical assistance, and education programs."

FDIC proposal providing a framework that centralizes and standardizes certain model risk management and third-party relationship due diligence functions through a voluntary certification process, overseen by regulators, could over time, significantly reduce the barriers to adoption of certified models by individual community banks and smaller institutions, thereby increasing the speed of adoption of innovative technology via well-vetted partnerships.[40]

Question 10: Please describe any particular challenges or impediments financial institutions face in using AI developed or provided by third partiesand a description of how financial institutions manage the associated risks. Please provide detail on any challenges or impediments. How do those challenges or impediments vary by financial institution size and complexity?

It is critical that regulators and bank supervisors acknowledge that partnerships with technology firms are likely *the only way* that the vast majority of banks and credit unions will be able to overcome the many barriers that stand in the way of a successful digital transformation of their traditional branch-based consumer lending programs. It is simply too much to expect that any but the largest banks in the United States will be able to organically develop the software, methods and the associated technical expertise, to manage a successful online consumer lending program that uses advanced AI/ML techniques. A successful program requires more than an "Apply Here" button on a bank's website. From online customer acquisition to fraud protection to underwriting and pricing, to meeting the demands of modern mobile and online experiences, to digital servicing and collections, in the current age, it is a complex enterprise.

Each of the prudential regulators has issued safety and soundness guidance for the institutions it supervises on managing risk in connection with the use of third-party vendors.[41] As has been recognized in this regulatory guidance, extensive oversight of third-party vendors is expected from financial institutions who use them. However, the specific application of this guidance to an institution's use of a third party's AI models, especially an activity that is deemed critical or high risk such as credit underwriting, is still uncertain. In particular, the third-party guidance does not specifically address how institutions, especially smaller institutions, are expected to manage risks presented by the use of vendor-provided tools -- such as AI-driven credit underwriting models -- that

[40]In this regard, last year the FDIC issued a request for information on standard setting and voluntary certification for models and third-party services providers, overseen by regulators *See* https://www.fdic.gov/news/press-releases/2020/pr20083a.pdf; 85 Fed. Reg. 44890 (July 24, 2020).

[41] See Board of Governors of the Federal Reserve System, SR Letter 13-19, "Guidance on Managing Outsourcing Risk" (December 5, 2013); Federal Deposit Insurance Corp., FIL 44-2008, "Third-Party Risk: Guidance for Managing Third-Party Risk" (June 6, 2008); National Credit Union Administration, SL No. 07-01, "Evaluating Third-Party Relationships" (October 2007); Office of the Comptroller of the Currency, Bulletin 2013-29, "Third-Party Relationships" (October 30, 2013). *See also* Consumer Financial Protection Bureau, Bulletin 2012-03, "Service Providers" (2012); OCC, Bulletin 2020-10, "Third-Party Relationships: Frequently Asked Questions to Supplement OCC Bulletin 2013-29" (March 5, 2020); FDIC, FI-50-2016, "Request for Comment on Proposed Guidance for Third-Party Lending."

have demonstrated value but operate at a level of sophistication that the institution's human and financial resources do not permit it to reproduce in-house. The small and mid-sized institutions are reluctant to commit resources to vendor relationships without assurances from their regulators about how they can do so in a manner that is consistent with the regulators' expectations for prudent risk management. Current guidance directed more specifically to the management of model risk, which pre-dates the agencies' third-party risk guidance, does not resolve this uncertainty.[42]

First, there are often varying interpretations of the existing model risk management governance of these models and the exact oversight responsibilities banks have when they engage a vendor that employs an AI model. Uncertainty surrounding the appropriate method for applying the existing model risk management guidance to third party AI technology -- and the supervisory application of the principles -- can discourage banks of any size from using AI-driven models. Note that banks can also be discouraged due to uncertainty about the appropriate fair lending testing regime that regulators expect will be applied.

The current model risk management guidance dates back 10 years to 2011, a time prior to the growth of modern AI applications. The guidance is not targeted specifically to the management of modern AI models or those developed and managed by third parties.[43] Work done by the Bank Policy Institute ("BPI") suggests that although the guidance technically "gives banks flexibility to modify the model risk management framework for validating vendor and other third party models," the reporting on the ground reveals that federal banking regulators "have not consistently afforded this flexibility to banks with regard to vendor-developed AI credit underwriting systems." According to BPI, regulators "have not applied a similar review or approval process to widely used conventional underwriting systems."[44] If this approach persists, it will create an unlevel playing field -- one that fails to harness the benefits of new models or acknowledge that traditional models may be less accurate and more biased against protected groups.

Second, absent new or revised formal guidance, many financial institutions may not be able to participate in the growing adoption of AI and may not become aware of the growing recognition of responsible AI/ML model use by federal regulators focused on both prudential supervision and regulation and on consumer protection. Federal regulators, therefore, have the opportunity to significantly improve banks' ability to use AI technology in the near term by issuing examiner guidance that would clarify the application of existing model risk management and third-party relationship risk

---

[42] Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency, "Supervisory Guidance on Model Risk Management" (2011); FDIC, FIL-22-2017, "Adoption of Supervisory Guidance on Model Risk Management" (June 7, 2017).
[43] Bank Policy Institute and Covington, "Artificial Intelligence Discussion Draft: The Future of Credit Underwriting: Artificial Intelligence and Its Role in Consumer Credit" (2019) at p. 6.
[44] *Id.*

management principles to AI/ML credit underwriting models, including those sourced from third-party vendors.

We recognize that recalibrating the approach to model risk management on an interagency basis is a difficult undertaking that may take significant time and effort. There are steps the agencies can take sooner rather than later, however, that would address the current uncertainty and facilitate the use of responsible innovation while that process is ongoing. For example, issuing examiner guidance in the near term would be consistent with regulators' longer-term effort to modernize its digital activities regulations. The guidance would not displace the current Model Risk Management guidance or the Third-Party Service Provider guidance, which would remain applicable. It would complement those documents in a relevant and practical manner and could be updated as principles-based guidance or regulations evolve.

Key concepts for any updated examiner guidance could include:

- Federal regulators should confirm that banks' reliance on independently validated AI credit underwriting models managed by third parties is recognized / appropriate;
- The updated guidance should revise existing methods used and recommended by bank examiners for validating and fairness testing, so they are effective and relevant for evaluating a complex AI model, meaning that validation activities would be conducted largely at the model level, rather than the variable level.
- The guidance should make clear that while examiners should expect banks to develop a "detailed knowledge" of vendor-provided models, "detailed knowledge" does not require banks to have a detailed understanding of the model at the code level, just as financial institutions are not currently required to understand third-party proprietary credit scoring models at the code level.
- In an effective program, "detailed knowledge" means an understanding of the different categories of variables, the techniques used by the model, and the key metrics by which model outputs are measured - providing banks with the ability to confirm that the use of the model is consistent with its prudent operation, safety and soundness, and fair lending. This is best accomplished by using appropriate metrics and regular tests for model accuracy and fairness.
- Finally, regulators' approaches to supervision should reflect that the practical application of model risk management principles to vendor-provided AI models will be different than it is in cases of bank-developed models.

Question 11: What techniques are available to facilitate or evaluate the compliance of AI-based credit determination approaches with fair lending laws or mitigate risks of non-compliance? Please explain these techniques and their objectives, limitations of those techniques, and how those techniques relate to fair lending legal requirements.

Applying AI and richer data sets to lending has great potential to make lending more fair and more inclusive than the current traditional system.[45] In general, technological advances have benefitted consumers seeking credit by reducing the scope of human bias and enhancing the reach of human intelligence. While imperfect, automated credit scoring opened up credit to individuals who may have lacked the kind of personal history or relationships formerly needed to apply for credit at a local bank. In the same way, AI models open up credit to individuals who lack the kind of credit file formerly needed to secure an appropriately robust credit score.

Still, we must also be clear-eyed about the potential risks and take steps to address them.[46] Models that employ facially-neutral criteria and operate on large volumes of data could still end up doing little to improve on the legacy of discrimination, or even may exacerbate our credit system's deeply unequal status quo. Separately, the large number of data sources used by AI/ML algorithms could increase these risks if they are not selected with care and monitored. Upstart has demonstrated that these risks can be effectively mitigated in AI-based credit underwriting if variables are closely monitored for bias, in line with regulatory expectations.

Upstart has developed and routinely applied sophisticated fair lending tests, as well as access-to-credit tests, to all lending outcomes on its AI platform over the past seven years (covering nearly one million borrowers). To date, Upstart can report that financial institutions' use of Upstart's AI model has enabled a significant "expansion of credit access…across all tested race, ethnicity, and sex segments" and that the Upstart model does not introduce bias into financial institutions' credit decision process.[47]

Further, Upstart's fair lending reporting procedures ensure that its bank partners can validate that future versions of the model continue to be fair. To ensure sound, effective fairness testing, there are a number of techniques Upstart uses, and principles that Upstart adheres to, that could help guide best practices in fair lending evaluations of AI models. These techniques and principles may also help regulators evaluate whether additional guidance could be provided to market participants.

The current application of the Equal Credit Opportunity Act ("ECOA") and its implementing Regulation B ("Reg B") has produced a number of well-known approaches for financial institutions to analyze disparities in lending. Each approach has some strengths and weaknesses when applied to credit models of any type. However, a complete and optimal solution depends in part on the public policy objectives being pursued but likely *requires the use of several tests in concert.*

---

[45] *See, e.g.*, Richard Cordray, Director, CFPB, Alternative Data Field Hearing (Feb. 16, 2017), *available at* https://www.consumerfinance.gov/about-us/newsroom/prepared-remarks-cfpb-director-richard-cordray-alternative-data-field-hearing/.

[46] Klein, Aaron. "Reducing Bias in AI-Based Financial Services." *Brookings Institute*. 10 July 2020. https://www.brookings.edu/research/reducing-bias-in-ai-based-financial-services/. 28 June 2021.

[47] "An update on credit access and the Bureau's first No-Action Letter" https://www.consumerfinance.gov/about-us/blog/update-credit-access-and-no-action-letter/

One of the simplest, yet effective, approaches to fair lending testing is to compare credit decisions on different groups of borrowers (statistical parity). For example, a test can assess whether different demographic groups have the same approval rate for a loan. The testing standard could use either a preset ratio threshold or a statistical significance test.[48] This test, however, suffers from an obvious drawback; because of large underlying socio-economic disparities between groups, different groups of borrowers in practice have different loan outcomes, i.e., they default at different rates when given loans. Forcing approval rate equality across groups would lead to either over-approving borrowers who would default, or under-approving borrowers who would repay. Both scenarios are bad for borrowers and conflict with financial institutions' core business objectives.[49]

A second essential testing methodology that can help address the shortcomings of statistical parity is a calibration test. Financial institutions will want to assess the parity of a model's output for different protected classes, conditional on the complete set of factors affecting creditworthiness, i.e., the actual outcomes that show which borrowers repay their loans. This can be implemented by testing the following question: When a model predicts a particular risk, does that translate to the same actual default rate across groups? If a lender finds that one demographic group actually defaulted at a higher rate than another group at the same predicted risk, that would justifiably cause concerns about fairness. Calibration is objective and generalizable. It doesn't incentivize distortionary behavior or incline financial institutions to approve overly risky borrowers. And it has practical relevance to borrowers; it evaluates both approval and APR decisions. This makes calibration a sound approach from a fair lending perspective, and calibration tests are widely used in practice. There are drawbacks here as well, however.[50]These are just two of the eight different types of fairness tests and supplemental approaches

---

[48] This is known as statistical parity in the literature. *See, e.g.*, https://arxiv.org/pdf/1703.09207: Berk, Richard, et al. "Fairness in criminal justice risk assessments: The state of the art." Sociological Methods & Research (2018): 0049124118782533.

[49] Given these realities, this standard could push financial institutions to avoid marketing to certain disadvantaged groups of consumers, even if many would repay, to avoid higher defaults.

[50] Details of a chi-squared test for calibration proposed by VantageScore for credit scoring models: https://www.vantagescore.com/images/resources/FINAL-Statistical%20Bias%20WP18-Online.pdf. The shortcomings of calibration include the fact that financial institutions don't observe outcomes at every level of predicted risk (they do not lend to borrowers whose predicted risk is higher than the risk tolerance). This is an effective test in the spectrum of risk scores where a lender lends, so the wider that spectrum is, the more complete this test is. Unfortunately, in some specific cases a model could be calibrated without being useful to consumers. Consider a hypothetical situation where we have a model based only on credit score, and two demographic groups. Suppose the credit score is very predictive for one demographic group, and therefore gives them a range of scores across the spectrum, but the score is not predictive for another group, for example because they lack credit history. A calibrated model could output a range of (cont.) predicted default probabilities for the first group, but would output the same (mean) score for everyone in the second group because they are indistinguishable from the perspective of the model. This would lead to a situation where the model approves nobody from the second group if their mean risk is above our risk tolerance. That would be unfair to applicants from the second group, some of whom at least would actually repay their loan. These concerns can be mitigated by specifically assessing their relevance to the lender in question, or by combining calibration with other fair lending tests.

that Upstart and other industry participants can apply to evaluate the fairness of an AI model.[51]

Upstart encourages the adoption of guidance on appropriate testing methods and basic principles that should be followed in fair lending testing. An effective fair lending testing regime must follow certain basic principles. First, testing should be objective rather than subjective, and the tests themselves should not leave room for human interpretation of whether a practice seems fair or reasonable. Furthermore, testing should be readily understandable and verifiable -- a lender should not be able to manipulate or hide the true results of the test.

A second key principle is that fair lending tests should be universal and generalizable. This means that tests should be applicable to all competing approaches to both underwriting and pricing. Upstart's thesis is that an effective test should not *a priori* assume that any specific approach to lending is fair, even if, for example, it has been used historically. Furthermore, the test of fairness, at least conceptually, should be extensible to different technical approaches and innovations, including both traditional methods of underwriting consumers and any new or innovative technologies that might become available.

Third, the most effective and appropriate fair lending tests measure quantities that are relevant to the consumer with achievable target thresholds. Tests should be designed to measure impacts on the consumer, rather than operating as a purely theoretical or intellectual exercise. The more complete a test's coverage of the quantities relevant to the end consumer, the more relevant the test becomes. For example, tests should measure key issues like approval rates and interest rates/APRs. Fair lending testing standards should actually be achievable in practice and take into account real world constraints. For example, certain disparity thresholds in lending may be nearly impossible to achieve in practice by any individual lender because of the large systemic inequalities in socioeconomic conditions outside of the lender's control. Therefore, for any fair lending test to not be self-defeating, it must be possible for a lender to meet that test sustainably and without undermining its business because a lender that stops lending responsibly does not serve the credit needs of consumers.

---

[51] These eight tests are (i) demographic parity test, (ii) constant test, (iii) classification parity test, (iv) calibration test, as well as supplemental tests such as (v) equal accuracy tests, (vi) comparison test, (vii) debias test, and (viii) tradeoffs test. See: "Does Credit Scoring Produce a Disparate Impact?" Federal Reserve Finance and Economics Discussion Series. Divisions of Research & Statistics and Monetary Affairs Avery, Brevoort, Canner.. "Fairness in criminal justice risk assessments: The state of the art." Sociological Methods & Research (2018) Berk, Richard, et al. "Fairness through awareness." Dwork, Cynthia, et al. "On conditional parity as a notion of non-discrimination in machine learning." Ritov, Ya'acov, Yuekai Sun, and Ruofei Zhao. "The measure and mismeasure of fairness: A critical review of fair machine learning." (2018) Corbett-Davies, Sam, and Sharad Goel. "Inherent trade-offs in the fair determination of risk scores." Kleinberg, Jon, Sendhil Mullainathan, and Manish Raghavan. "Tracking and Improving Information in the Service of Fairness." Proceedings of the 2019 ACM Conference on Economics and Computation. (2019). Garg, Sumegha, Michael P. Kim, and Omer Reingold. "Certifying and removing disparate impact." Feldman, Michael, et al. ACM international conference on knowledge discovery and data mining (2015).

Fourth, Upstart's view is that effective fair lending tests should not inadvertently limit progress towards important fairness objectives, such as equitable access and financial inclusion, because the testing focus is limited to only certain aspects of fairness or seeks elimination of only certain types of disparity. Given the consensus that the status quo in credit scoring and access to credit is far from ideal, it is important for any testing regime not to lock-in the status quo. Among other things, this means a fairness test should not be structurally anti-change or incumbent-preferring.

Finally, fair lending standards should not encourage financial institutions to make loans to borrowers who will likely be unable to repay them. Extending consumer credit is not beneficial to all consumers in all circumstances. Any fair lending standards that compel a lender to extend credit to a consumer who is unlikely to be able to repay the loan may lead to default, bankruptcy or financial hardship. Fair lending tests should avoid inadvertently distorting lender incentives towards practices that would run counter to key policy goals, such as access; lending responsibly, i.e., to those with ability to repay); and fairness, including in approvals and pricing.

Question 12: What are the risks that AI can be biased and/or result in discrimination on prohibited bases? Are there effective ways to reduce risk of discrimination, whether during development, validation, revision, and/or use? What are some of the barriers to or limitations of those methods?

Too often, human intelligence has proven no match for human biases. The triumph of such biases over human intelligence has often resulted in systemic discrimination. AI offers the potential to apply intelligence to credit decisions without bias. In this way, AI is clearly part of the solution, rather than the problem, addressed by discrimination law. The possibility that AI models may reflect pre-existing human bias is not a persuasive reason to avoid AI in favor of existing models that may also be infected by human bias.

Although an AI system itself may not be intrinsically biased, as noted above, AI model outputs may nonetheless reflect or reinforce the underlying discrimination and disparity in society. Upstart has empirically demonstrated for several years that consistent implementation of robust monitoring and controls can effectively help mitigate the risk of bias in AI. Upstart has found that there are effective ways to reduce the risk of having the technology simply reinforce existing biases, discrimination, and inequality in society, such as through rigorous testing and monitoring.

Upstart also firmly believes that oversight, transparency and diverse perspectives are important in reducing the risks associated with AI. In September 2017, Upstart became the first company to apply for and receive a No-Action Letter from the Consumer Financial Protection Bureau (CFPB).[52] The purpose of such letters is to reduce potential regulatory uncertainty for innovative products that may

[52] See https://www.consumerfinance.gov/about-us/newsroom/cfpb-announces-first-no-action-letter-upstart-network/

offer significant consumer benefit. On November 30, 2020, at the expiration of the first No-Action Letter, Upstart received a new No-Action Letter from the CFPB, which has a three-year term.[53] A component of these No-Action Letters requires Upstart to provide periodic reporting and other information relating to its AI model development and fair lending testing results and analyses to the CFPB. This provides the CFPB with both oversight of, and insight into, the methods and techniques that may be deployed in the real world to minimize the risks associated with AI lending models. In addition to regulators, Upstart also works with other independent third-party stakeholders and experts to evaluate model performance on fairness grounds.[54]

While a model is in development, it is critical for model developers to eliminate any variables that directly identify protected classes as well as variables that may be close proxies or substitutes for those variables. Quantitative techniques exist to assess whether a predictive variable, either independently or in conjunction with other variables, may be acting as a close proxy for a protected group. These techniques could be even more effective in eliminating biases in the model if proxy standards were more clearly defined, as discussed in more detail below.[55]

Some commentators have proposed the possibility of conducting model training utilizing protected variables to attempt to reduce disparate impact in any model, including AI/ML models. Upstart is continually evaluating new approaches for reducing disparity and believes that such approaches may be effective in reducing these disparities. In Upstart's view, however, there is significant uncertainty under ECOA surrounding the legality of using protected variables (or proxy estimates of the same) in model training, even when such variables are used for the purpose of reducing disparate outcomes, and would welcome more guidance from regulators in this area.[56]

There are a number of other regulatory barriers that impede the adoption of various risk-mitigation techniques. The industry currently lacks consistently applied standards for identifying prohibited discrimination and reducing bias in AI. Two specific areas where regulators could improve clarity include proxy methodology and variable selection. With respect to proxy methodology, existing methodologies differ by regulator and are imperfect in application, particularly as applied to new

---

[53] See https://www.consumerfinance.gov/about-us/newsroom/consumer-financial-protection-bureau-issues-no-action-letter-facilitate-use-artificial-intelligence-pricing-and-underwriting-loans/

[54] In December 2020 NAACP Legal Defense and Educational Fund, Inc. and the Student Borrower Protection Center announced a collaboration with Upstart to review of Upstart's fair lending outcomes and assess best practices in the use and testing of alternative data in fintech credit models, see https://www.naacpldf.org/press-release/naacp-legal-defense-and-educational-fund-and-student-borrower-protection-center-announce-fair-lending-testing-agreement-with-upstart-network/

[55] See Kallus, Nathan & Mao, Xiaojie & Zhou, Angela. (2019). Assessing Algorithmic Fairness with Unobserved Protected Class Using Data Combination.

[56] Upstart to-date has generally been discouraged by regulators from undertaking this type of exploration due to the associated uncertainty. See: "What's in a Name? Reducing Bias in Bios without Access to Protected Attributes" https://www.aclweb.org/anthology/N19-1424.pdf

technology and/or model advancements.[57] The industry and consumers would benefit from the adoption of an updated and uniformly accepted proxying standard to use when conducting fair lending analysis to ensure that there is consistency in the industry.

As to variable selection, there are presently many diverse sets of data available to creditors that were not available when ECOA was enacted. While these new data sets may carry the potential for bias if used inappropriately, they also carry a tremendous opportunity for consumers to benefit by improving accuracy that may in turn enable better-priced credit. These benefits cannot be realized if their use is outright prohibited or deemed too risky to explore. Accordingly, regulators should avoid a focus on identifying prohibited variables and instead focus on guidance for establishing robust standards for data integrity and fair lending testing methods to combat bias.

Question 13: To what extent do model risk management principles and practices aid or inhibit evaluations of AI-based credit determination approaches for compliance with fair lending laws?

As a service provider subject to the BSCA and the consumer protection laws that apply to all financial institutions, Upstart relies upon the existing model risk management guardrails for its model risk program, which consists of a number of systematic and operational procedures designed to reduce risk by providing reasonable assurance the model is operating as intended, ensuring ongoing model improvements to maintain effectiveness, and promoting effective oversight through strong model controls and validation.

The existing guidance issued by the Federal Reserve, OCC, and FDIC for model risk management is not targeted specifically to AI models and thus has gaps with respect to the appropriate model risk management governance around those models and what oversight responsibilities banks have when they engage a vendor that uses AI models. . When it was issued, the guidance was a forward-leaning and essential effort by the agencies to set guardrails for institutions' use of models in a variety of areas. The guidance has since been overtaken in important ways -- including the increasing use of AI and ML in model development -- in the 10 years since the agencies released it. While much of the framework can be retained, the agencies should tackle revising the guidance specifically with the goal of facilitating the use of AI/ML-driven models, including by community institutions. That would include addressing at minimum, the following considerations:

- To what extent must a bank understand how a complex AI/ML credit underwriting model operates in order to meet its oversight obligations? The principles in the existing model risk management guidance should be updated so that they specifically reference AI model

[57] In Upstart's experience, financial institutions are not uniformly familiar with the BISG methodology and / or may assert that their primary regulator relies on a different standard / approach.

governance, and the existing methods used by the agency's examiners for validating and testing should be revised so they are effective for complex AI models. Validation activities should be conducted at the model level, not at the variable level.

- What level of detail is appropriate for model documentation provided by third-party vendors? Examiners should expect financial institutions to develop appropriately detailed knowledge of vendor-provided models. The knowledge required of a financial institution should not mean that the bank must understand the model at code level. Instead, the financial institution should be expected to have sufficient knowledge to confirm that the use of the model is consistent with its prudent operation and safety and soundness with an emphasis on using appropriate metrics and tests for model accuracy. The agencies should clarify specifically how financial institutions should satisfy that expectation.

- Most banks will need to rely on external expertise in AI lending because of constraints on their resources. The agencies should make clear that such reliance on third parties is appropriate, and clarify to what extent may financial institutions may rely on third parties to audit and validate a vendor-provided AI/ML credit underwriting model as a component of the due diligence and ongoing oversight process. The agencies' approach to supervision should reflect that the practical application of model risk management principles to vendor-provided AI models will be different than to bank-developed models.

In Upstart's experience, financial institutions rely heavily on supervisory examiner communications to determine what oversight and model governance standards should apply. However, these communications at times may be inconsistent with official agency policy, regulation, or guidance and/or also inconsistent with instructions or guidelines provided by other federal regulatory agencies. If the consumer benefits of using alternative data and innovative modeling techniques are to be fully realized in underwriting through widespread adoption, the various risk-management guidelines must be clarified, on an interagency basis, or alternatively, by a single agency given authority to prescribe the necessary uniform guardrails regarding oversight of these new technologies.

Fair lending compliance is well understood to be an essential component of the model oversight responsibilities for any lender using an AI underwriting model, but there currently is substantial uncertainty about what is required. Issues for the agencies to consider include:

- How best to achieve a single set of standards and testing requirements that could be consistently applied, including a single approach to proxy methodology;
- How best to develop testing using disparate impact theories to ensure that AI models are not introducing discrimination into the underwriting process;
- How best to encourage the use of alternative variables in underwriting so long as variables are introduced in a responsible manner with appropriate accuracy and fair lending testing.

It is also critical to note historically the agencies have expected banks to keep model development and testing separate from fair lending testing. In addition, older model risk management guidelines that do not reflect the unique characteristics of AI/ML models may discourage banks from using AI/ML credit underwriting models. These factors may also, in turn, limit financial institutions' ability to reduce disparities that exist in their consumer lending. A more coordinated regulatory approach that better integrates prudential and consumer protection considerations for model development and validation testing and that better reflects the nature of AI/ML models would benefit consumers, industry, and regulators alike.

For example, a key pillar of an effective fair lending assessment is reviewing whether there are less discriminatory alternatives than the one currently in use. Today, banks are likely to find themselves with traditional credit models that produce significant disparity or discrimination. Many still use blunt instruments such as high credit score cut-offs that disproportionately hurt minority borrowers. Even where they are aware that they could produce a less discriminatory outcome if they adopt an AI credit model that harnesses alternative data, many banks are reluctant to do so and report informally that it is because of the uncertainty surrounding existing model risk management guidelines and particularly, uncertainty regarding the model documentation banks are required to maintain, and the level of technological sophistication they must demonstrate, regarding their oversight of a vendors' AI model.

Financial institutions contemplating using third-party credit models would also benefit from additional clarity on best practices for AI model documentation and validation. For model documentation, if an institution is able to understand and report on the techniques and variables used at a general conceptual level, they should feel comfortable working with the third-party model developer. With respect to model validation, both prior to its use and/or as part of an institution's ongoing monitoring activities, an institution should be able to rely on third-party experts to assist in meeting such validation requirements, as long as they understand at a general level the validation procedures.

Question 14: As part of their compliance management systems, financial institutions may conduct fair lending risk assessments by using models designed to evaluate fair lending risks ("fair lending risk assessment models"). What challenges, if any, do financial institutions face when applying internal model risk management principles and practices to the development, validation, or use of fair lending risk assessment models based on AI?

Financial institutions should continue to be able to rely on their technology partners to help them develop and operationalize rigorous "fair lending risk assessment models" that use well-established,

regulator-approved testing techniques.[58] Another potential challenge for financial institutions in this area could be the expectation on the part of some bank supervisors that banks' internal model risk management principles and practices and compliance management systems must address in detail the technical aspects of any partnerships between a financial institution and third-party model developer, including the institution's use of a third party's AI models.

While fair lending examination guidance has been issued by the prudential regulators, that guidance is now somewhat dated, and there are opportunities for further clarity.[59] For example, guidance would be beneficial as to what testing outcomes require additional actions. Reg B and its commentary allow a credit practice that "has a disproportionately negative impact on a prohibited basis" provided that a lender can demonstrate "a legitimate business need that cannot be reasonably achieved as well by means that are less disparate in their impact."[60]

Question 15: The Equal Credit Opportunity Act (ECOA), which is implemented by Regulation B, requires creditors to notify an applicant of the principal reasons for taking adverse action for credit or to provide an applicant a disclosure of the right to request those reasons. What approaches can be used to identify the reasons for taking adverse action on a credit application, when AI is employed? Does Regulation B provide sufficient clarity for the statement of reasons for adverse action when AI is used? If not, please describe in detail any opportunities for clarity.

ECOA and Reg B require creditors to provide credit applicants with a notice of action taken within 30 days after receiving a completed application, and the reasons provided to the applicant for any adverse action must accurately represent the actual factors that were used in the decision on the application. Prospective borrowers who are unable to obtain credit on the terms offered by the financial institution must have the opportunity to learn the primary reasons why their application is denied.

As discussed in our response to Question 1, given the unique attributes of AI credit underwriting systems, generating adverse action reasons can require methods that are quite different from those employed to explain decisions made by traditional models.[61] And as outlined in our response to Question 1, there is a growing body of well-established techniques that can facilitate interpretation of complex AI and ML models (e.g., partial dependence plots, relative importance, Shapley values). These techniques offer financial institutions ways to quantify the impact of particular data sets and even

[58] In Upstart's experience, applying measures of fairness (tests outlined in our response to question 11) do not require harnessing AI/ML technology, but seeking alternatives as required under ECOA Reg B may require its use.
[59] See Federal Financial Institutions Examination Council, Interagency Fair Lending Examination Procedures (Aug. 2009), https://www.ffiec.gov/PDF/fairlend.pdf; Policy Statement on Discrimination in Lending, 59 Fed. Reg. 18,266 (April 15, 1994), https://www.govinfo.gov/content/pkg/FR-1994-04-15/html/94-9214.htm.
[60] See Official Interpretation of §1002.6(a) "General rule concerning use of information."
[61] Bank Policy Institute and Covington, "Artificial Intelligence Discussion Draft: The Future of Credit Underwriting: Artificial Intelligence and Its Role in Consumer Credit" (2019) at p. 6.

individual variables (e.g, past loan delinquencies). These techniques can surface accurate explanations related to the impact of these individual variables or groups of related variables on the model's output for a particular individual.

Upstart has demonstrated that the existing regulatory framework has sufficient flexibility to adapt to AI/ML where the variables and key reasons for a denial may be less intuitive.[62] ECOA and Reg B do not require customized descriptions of how or why a particular factor adversely impacted each individual applicant. Furthermore, neither ECOA nor Reg. B mandates the use of a specific list of decline reasons or a particular methodology for selecting such decline reasons.[63] This provides general flexibility for the deployment of AI-based AAN models and corresponding denial notices. Unfortunately, the existing sample notification forms in Appendix C of Reg B do not reflect AI/ML advancements. Rather, the ten model forms are all focused on traditional underwriting criteria. Therefore, in the spirit of furthering innovation and ensuring accurate and informative communications to the consumer, it could be beneficial for the CFPB to provide guiding principles for determining and disclosing decline reasons in AAN disclosures and consider updating sample notification forms to account for various types of credit modeling techniques, including AI.

Because AI models often employ larger datasets than traditional credit models, individual variables become less heavily relied upon by the model and therefore may be less influential in a credit denial. In some cases there may be many more than 4 variables (or variable groupings) that influence the model's risk assessment of the application. Financial institutions should be encouraged to investigate the interactions between variables to produce accurate and informative candidate denial reasons that are not based on single variables but rather on groupings of closely correlated variables (see Shapley information above as one possible technique). Additionally, when there are multiple, equally significant reasons for denial and financial institutions must select which to display to an applicant, priority could be placed on disclosing reasons that are within the applicant's control to change. This approach would promote financial literacy and expansion of credit access.

Adverse action notice reasons is a critical topic for AI/ML credit underwriting models. It is important to note that Upstart has found that it is possible to search for, surface and present meaningful explanations to applicants from its AI model, with a standard format. In general, Upstart believes that the explanations should be provided to applicants and other users in a digestible form that is rank-ordered in a way that is relevant and actionable. For instance, it would not be useful for a system to explain itself so comprehensively -- including all the relevant explanations provided for every step taken by a complex system -- that the impacted applicant or other user cannot understand what is actionable or not actionable on their part. Rather, on the Upstart platform, its adverse action system determines which rejection reasons are most applicable to the applicant by evaluating the individual

[62] 12 C.F.R. pt. 1002, comment 9(b)(2).
[63] *Id.*

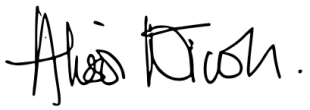
variables (or clusters of highly correlated variables) that were of most significance in the underwriting model's assessment of that particular application. Denied applicants are additionally provided with a copy of the information submitted in the application which was used to inform the decision as well as the requisite disclosures regarding instructions on how to obtain a copy of the consumer report that was used in connection with the application, where applicable. Under such an approach, applicants have the information to take whatever action may be necessary to improve their credit profile or correct any credit file inaccuracies.

Question 16: To the extent not already discussed, please identify any additional uses of AI by financial institutions and any risk management challenges or other factors that may impede adoption and use of AI.

Question 17: To the extent not already discussed, please identify any benefits or risks to financial institutions' customers or prospective customers from the use of AI by those financial institutions. Please provide any suggestions on how to maximize benefits or address any identified risks.

Upstart thanks each of the agencies for its ongoing dialogue and for the opportunity to comment on the RFI. Upstart looks forward to continuing its ongoing dialogue with each agency on the issues raised in the RFI and other technology-related issues. Upstart appreciates the agencies' focus on technology issues in the banking industry, including through the use of AI in financial services. If you have any questions, please contact the undersigned by phone at (833) 212-2461 or by email at alison@upstart.com.

Sincerely



Alison Nicoll
General Counsel
Upstart Network, Inc.

cc: Paul Gu, SVP Product & Data Science, paul@upstart.com
Nathaniel Hoopes, VP, Head of Government Relations and Regulatory Affairs, nat.hoopes@upstart.com