



August 21, 2023

RE: Quality Control Standards for Automated Valuation Models:

Docket ID OCC–2023–0002; FDIC RIN 3064–AE68; Federal Reserve Docket No. R–1807 and RIN No. 7100 AG60; Docket Number NCUA–2023–0019; FHFA RIN 2590–AA62; and CFPB–2023–0025

To Whom it May Concern:

Thank you for the opportunity to comment on the use of automated valuation models (AVMs)¹ by mortgage originators and secondary market issuers.

We believe that the Office of the Comptroller of the Currency (OCC), the Federal Deposit Insurance Corporation (FDIC), the Federal Reserve (FRB), the National Credit Union Administration (NCUA), the Federal Housing Finance Agency (FHFA), and the Consumer Financial Protection Bureau (CFPB) (collective, the "Agencies") must address standards for the use of AVMs.² However, with advancements in artificial intelligence, new capabilities to source more data, the influence of online real estate platforms that use them, as well as receptivity at the Federal Housing Finance Administration for their use in evaluating portfolios,³ the reliance on AVMs by actors in real estate markets will increase.

While the thrust of the proposed rule focuses on addressing an observed bias toward overvaluation, it is also essential that the Agencies pay close attention to the possibility that AVMs could perpetuate biases that undermine protected class members.

The proposed rule says that using AVMs can be considered a sound banking practice but includes several controls to explain how supervised banks can conform to meet standards for safety and soundness. Implicitly, the Agencies have stated that AVMs are appropriate for assessing collateral in mortgage

¹ Section 1125 (amended title XI of FIRREA) defines an AVM as “any computerized model used by mortgage originators and secondary market issuers to determine the collateral worth of a mortgage secured by a consumer’s principal dwelling.”

² Porter, T. J. (2023, June 2). What Is An Automated Valuation Model (AVM)? *BankRate Real Estate*. <https://www.bankrate.com/real-estate/automated-valuation-model-avm/>

³ Mortgage Bankers Association. (2019). *The State of Automated Valuation Models in the Age of Big Data* (Valuation Analytics Working Group Paper) [Member White Paper]. <https://www.mba.org/docs/default-source/uploadedfiles/member-white-papers/stateofautomatedvaluationmodels-final>

origination or operation of secondary markets. We recognize the wisdom of this approach. It allows banks to take advantage of technological advancements but makes clear that their use will not permit evasions of pre-existing safeguards.

SUMMARY

I. AVMs, if supervised correctly, can have beneficial effects. However, they must be accurate to preserve financial stability in housing markets.

i. The quality control for high levels of confidence in accuracy (the “accuracy” control) should assess variations from contract prices. It should be sensitive to gaps that undervalue and overvalue properties.

ii. Because AVMs can make analyses in near real-time, they could make markets fairer, particularly for lower-wealth households.

iii. AVMs can ensure that secondary market issuers adequately describe the value of the collateral used in securitization. The Agencies will support financial stability by applying quality controls to secondary market issuers.

II. Coverage should include all AVMs used to value dwellings, including ones that involve owner financing

i. The agencies should clarify that covered mortgages include loans for residential dwellings not secured by real property and installment loans.

ii. The definition of mortgage should include principal dwellings secured by a purchase money installment contract (owner financing).

III. The Agencies must ensure that AVMs do not have discriminatory impacts. As a part of doing so, they should implement standards for testing models.

i. Research demonstrating how appraisal bias has contributed to discrimination in housing markets underscores why the Agencies must include a fifth quality control to protect against discriminatory models.

ii. The Agencies should establish a fifth quality control to ensure that AVMs are held accountable if they are discriminatory.

iii. Applying a fifth control against discrimination will reinforce essential aspects of the other quality controls.

iv. While covering a different activity, The PAVE Report and Freddie Mac's research on appraisals can still provide insight into how Agencies might measure discriminatory bias in AVMs.

v. AVMs, like all algorithmic processes, can achieve dual goals of accuracy and fairness.

vi. Creators of AVMs would benefit from clarity on how they can self-test their models for discrimination.

vii. Agencies should release data to the public to allow third parties to review AVM models for fairness.

viii. Tests based on outcomes rather than just inputs will be valuable when reviewers determine if an AVM is discriminatory.

iv. Reviews should ensure that models are trained on representative data sets. Reviews can use statistical procedures to determine which inputs contribute to discrimination.

x: Appraisals made by an AVM should be explainable.

DISCUSSION

I. AVMs, if supervised correctly, can have beneficial effects. However, they must be accurate to preserve financial stability in housing markets.

i. *The quality control for high levels of confidence in accuracy (the “accuracy” control) should assess variations from contract prices. It should be sensitive to gaps that undervalue and overvalue properties.*

Several empirical studies have determined that up to 90 percent of appraisals overestimate collateral value. Some research suggests that inaccuracies may be especially profound and artificially higher for assessing the value of homes in rural areas,⁴ distressed sales or for unique properties.⁵ Lax underwriting standards contributed to the financial crisis.⁶

The Agencies have rightfully planned to create a quality control to “ensure a high level of confidence in the estimates produced by AVMs.”

Incorrect valuations can lead to adverse consequences on either side of the contract price. Lower valuations may reduce access to credit, leading to adverse decisions that prevent loans from being approved and undermining the attainment of homeownership. However, valuations that are too high present risks that borrowers will not be able to repay a loan. The possibility exists that an overly-high valuation – both with a home purchase or a refinance – could mean that borrowers lose equity.⁷

Additionally, exaggerated estimates can have consequences for systemic financial stability. Overvaluation creates systemic risks that put too many borrowers at risk of being overloaded with debts they cannot afford. Researchers have concluded that inflated appraisals contributed to the financial crisis and that inflated estimates were disproportionately more likely to happen on homes purchased by lower-income borrowers. Unsurprisingly, loans with higher valuations had higher prices, and borrowers had higher debt loads. Borrowers in these situations were more likely to default.⁸

Valuations that are too high may permit exaggerated home sale prices, leading to unrealistic debt loads for borrowers and potentially fueling home price inflation. Valuations that are too low, while not creating

⁴ Bogin, A. N., & Shui, J. (2020). Appraisal Accuracy and Automated Valuation Models in Rural Areas. *The Journal of Real Estate Finance and Economics*, 60(1), 40–52. <https://doi.org/10.1007/s11146-019-09712-0>

⁵ Mortgage Bankers Association. (2019). *The State of Automated Valuation Models in the Age of Big Data* (Valuation Analytics Working Group Paper) [Member White Paper]. <https://www.mba.org/docs/default-source/uploadedfiles/member-white-papers/stateofautomatedvaluationmodels-final>

⁶ Phil Angelides & Bill Thomas. (2011). *The Financial Crisis Inquiry Report: Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*. Financial Crisis Inquiry Commission. <https://www.govinfo.gov/content/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf>

⁷ Maureen Yap, Morgan Williams, Lisa Rice, Scott Chang, Peter Christensen, & Stephen M. Dane. (2022). *Identifying Bias and Barriers: Promoting Equity*. National Fair Housing Alliance. https://nationalfairhousing.org/wp-content/uploads/2022/02/2022-01-28-NFHA-et-al_Analysis-of-Appraisal-Standards-and-Appraiser-Criteria_FINAL.pdf

⁸ Ben-David, I. (2011). Financial Constraints and Inflated Home Prices during the Real Estate Boom. *American Economic Journal: Applied Economics*, 3(3), 55–87. <https://doi.org/10.1257/app.3.3.55>

systemic risk on housing price stability, can compromise the vital policy goal of advancing homeownership.

As a result, the Agencies must apply a quality control standard to ensure the accuracy of AVMs. The framework for measuring the impacts of valuations must consider problems that occur when results are too low and too high.

ii. Because AVMs can make analyses in near real-time, they could make markets fairer, particularly for lower-wealth households.

In practice, the length of time needed to receive a completed appraisal undermines the ability of borrowers to compete with cash buyers. One large mortgage lender tells borrowers to expect to wait one to two weeks to receive the results of an appraisal.⁹ Research suggests that the time-to-close is another contributor to bias that disadvantages households of color. When NCRC conducted a mystery-shopper test where sets of appraisers were asked to value seven homes in Maryland, using the same home but with pairs of Black and White applicants, one appraiser refused to complete an assignment, and another failed to complete the work.¹⁰

The duration to complete an appraisal has deleterious consequences for economic inequality. Realtors acknowledge that cash buyers have a significant advantage over borrowers because they can close sooner.¹¹ Large “iBuyers” use the quickness of cash to incentivize sellers to take lower offers.¹² In 2023, boomers will buy more homes than millennials, primarily because they can convert their existing home equity into cash and then offer to close faster than the millennials, who have to qualify for a loan.¹³ In 2022, only 3 percent of first-time homeowners paid with cash, compared to 27 percent of buyers who already owned a home.¹⁴ These factors reveal how the time needed to receive an appraisal undermines the pro-social goals of homeownership, reducing racial wealth inequality, and building wealth.

Responding to that problem, some borrowers may attempt to catch up by waiving an inspection. That is dangerous.

AVMs' speed and lower cost raise the possibility that AVMs can improve markets. They can level the playing field between cash buyers and borrowers. Their use may lower borrowing costs in other lending contexts, such as home equity lines of credit (HELOCs). However, for AVMs to truly gain acceptance in the minds of lenders and secondary market issuers, they must demonstrate their accuracy. The argument

⁹ Victoria Araj. (2023, June 23). How Long Does An Appraisal Take? | Rocket Mortgage. *Rocket Mortgage Learn*. <https://www.rocketmortgage.com/learn/how-long-does-an-appraisal-take>

¹⁰ Jake Lilien. (2022). *Faulty Foundations: Mystery-Shopper Testing In Home Appraisals Exposes Racial Bias Undermining Black Wealth*. National Community Reinvestment Coalition. <https://ncrc.org/faulty-foundations-mystery-shopper-testing-in-home-appraisals-exposes-racial-bias-undermining-black-wealth/>

¹¹ Jacqueline DeMarco. (2023, February 28). *Why a Cash Offer Beats a Financed One When Home Buying—Orchard*. <https://orchard.com/blog/posts/why-are-cash-offers-better>

¹² Ryan Fink. (2022, July 26). Why Is a Cash Offer Better Than a Mortgage? - iBuyer Blog. *iBuyer Blog*. <https://ibuyer.com/blog/cash-offer-vs-mortgage/>

¹³ Julian Mark. (2023, August 12). Homebuying gets harder for millennials as boomers' market share grows. *Washington Post*. <https://www.washingtonpost.com/business/2023/08/12/millennial-home-buyers/>

¹⁴ Dr. Jessica Lautz. (2022). *Eleven Takeaways From the 2022 Profile of Home Buyers and Sellers* (Profile of Home Buyers and Sellers). <https://www.nar.realtor/blogs/economists-outlook/eleven-takeaways-from-the-2022-profile-of-home-buyers-and-sellers>

that the accuracy of any AVM-generated valuation is inherently constrained because no one visits the home in person seems well-founded. Nonetheless, as far back as 2019, research by industry has grown to view them favorably. One study by the Mortgage Bankers Association found AVMs were becoming much more accurate because of their ability to draw from more data and to provide statistical insights (estimate ranges, hit rates, reliability metrics) that appraisals do not.¹⁵ However, big data and generative AI could reduce those risks.

iii. AVMs can ensure that secondary market issuers adequately describe the value of the collateral used in securitization. The Agencies will support financial stability by applying quality controls to secondary market issuers.

The proposed rule would require secondary market issuers to indicate how mortgage originators can comply with quality control standards when they use an AVM to make an appraisal waiver decision.

We support the Agencies' intention to use this approach because it fits naturally with market behavior. Secondary market issuers significantly influence how mortgage originators perform their underwriting. Their power is especially true when originators use Desktop Underwriter.

Moreover, statistical tools used to measure portfolios, including loans that received an appraisal waiver, can inform secondary market issuers about portfolio risk. To the extent that uncertainty exists when using an AVM for an individual home, the risk is less across an entire portfolio. Neighborhood-level inputs, for example, could minimize the impact of a single property with a problem that is not readily visible to an AVM.

Using AVMs to assess collateral risk is analogous to how mark-to-market accounting may reveal short-term weaknesses in a bond portfolio. If a "secondary market issuer" (SMI) makes a covered "securitization determination" (SD) using an AVM, the evaluation of collateral must be accurate to protect investors.

Because doing so will lessen systemic risk, we support the intention of the rule to require the application of quality control standards when any secondary market issuer uses an AVM after making an appraisal waiver decision.

II. Coverage should include all AVMs used to value dwellings, including those that involve owner financing

i. The agencies should clarify that covered mortgages include loans for residential dwellings not secured by real property and installment loans.

The Agencies seek comments on how to define the term "mortgage" in the context of collateral estimations made by AVMs. Section 1125 does not define the term "mortgage." The Agencies should provide clarity on two aspects of the question.

We call on the Agencies to clarify that coverage extends to all cases where a loan would apply for a dwelling, even when the collateral is not secured by land or classified as real property. This approach ensures coverage includes transactions for manufactured housing classified as personal property and accessory dwelling units (ADUs). Approximately six percent of US households live in housing not

¹⁵ Mortgage Bankers Association. (2019). *The State of Automated Valuation Models in the Age of Big Data* (Valuation Analytics Working Group Paper) [Member White Paper]. <https://www.mba.org/docs/default-source/uploadedfiles/member-white-papers/stateofautomatedvaluationmodels-final>

classified as real property. ADUs are a relatively new element of housing stock, but efforts to zone for their use in residential neighborhoods implies how they will grow in popularity. Both examples deserve to be covered by this regulation. The Agencies should confirm (Question 22) that the meaning of the term "mortgage" is consistent with the definition of "residential mortgage transaction" in Regulation Z, which defines the term to mean a transaction associated with a loan where a security interest is placed against a consumer's principal dwelling.

ii. The definition of mortgage should include principal dwellings secured by a purchase money installment contract (owner financing).

Additionally, some consumers finance the purchase of their dwellings through installment sales contracts. Because the rule intends to provide safeguards for AVMs used to assess dwellings, the definition of a covered mortgage should be expanded to include extensions of credit for property that will be used as a dwelling and purchased by an installment loan. Because purchase money installment contracts are analogous to "owner financing," expanding the term "mortgage" to include this category will broaden the rule to cover a segment of the market primarily composed of underserved borrowers. Applying quality controls for AVMs used in these contracts will provide consumer protection in a space where consumers are often vulnerable to coercive agreements.

This expansive definition will provide safeguards in manufactured housing communities (MHCs). It is not unusual for a resident to purchase a home from a park owner through an owner-financing arrangement. Typically, park owners acquire manufactured homes when a resident abandons them. Because many owners prefer not to own the land but not the houses, such a sale is a standard business practice. The buyers of these homes are usually people living at the margins who lack robust credit. By their nature, they are coercive transactions where a home's actual value is unknown. Given the information asymmetries in the transaction, and the likelihood that no appraisal will occur, an AVM would be a natural use case to provide certainty to borrowers at a minimal cost.

III. The Agencies must ensure that AVMs do not have discriminatory impacts. As a part of doing so, they should implement standards for testing models.

i. Research demonstrating how appraisal bias has contributed to discrimination in housing markets underscores why the Agencies must include a fifth quality control to protect against discriminatory models.

In 2022, the Interagency Task Force on Property Appraisal and Valuation Equity (PAVE) published the *Action Plan to Advance Property Appraisal and Valuation Equity*. It outlined how racially-driven biases in residential property valuation have contributed to racial wealth gaps and called for an interagency response.¹⁶

Research consistently reveals discrepancies in how AVMs impact how borrowers of color access credit. When NCRC conducted paired testing on identical properties, it found differences in appraisal values and

¹⁶ Interagency Task Force on Property Appraisal and Valuation Equity. (2022). *Action Plan to Advance Property Appraisal and Valuation Equity*. Biden-Harris Administration. <https://pave.hud.gov/actionplan>

treatment of applicants.¹⁷ A national sample of mortgage refinances from 2000 to 2007 found systematically lower appraisals for minority homeowners.¹⁸

ii. The Agencies should establish a fifth quality control to ensure that AVMs are held accountable for discriminatory impacts.

Discrimination is real, widespread, and longstanding. It harms communities of color by limiting housing access and thwarting wealth-building efforts.

The proposed rule requires mortgage originators and secondary market issuers to adopt and maintain control policies for four areas: to ensure confidence in the accuracy of estimates, to prevent data manipulation, to avoid conflicts of interest, and to implement random sample testing.

However, in the proposed rulemaking, the Agencies acknowledge that Section 1125 of Dodd-Frank permits the agencies to implement additional quality control standards for “any other such factor that the agencies determine to be appropriate.”

In the proposed rule, the Agencies have proposed to use their authority under Section 1125 to require mortgage originators and secondary market issuers to “adopt policies, practices, procedures, and control systems to ensure that AVMs used in connection with making credit decisions or covered securitization determinations adhere to quality control standards designed to comply with applicable nondiscrimination laws.”

We support the plans of the Agencies to require mortgage originators and secondary market issuers to be held accountable to meet non-discrimination laws. A control against discrimination must be stated explicitly to avoid evasive regulatory interpretations.

iii. Applying a fifth control against discrimination will reinforce essential aspects of the other quality controls.

Discriminatory models permit social bias to cloud rational estimates of collateral values. All forms of bias create unfounded judgments. Judgments made without any basis will lead to errors in risk assessments. An AVM that embeds discrimination into valuations will be less accurate because it will produce results below actual market values. Given how important it is for policies to protect the financial interests of all stakeholders, any process that introduces errors must be constrained.

Additionally, a control against discrimination supports essential needs in the market to prevent data manipulations. Models can be built on inappropriate data. The Agencies should articulate that AVMs using unrepresentative data sets qualify as data manipulation.

¹⁷ Jake Lilien. (2022). *Faulty Foundations: Mystery-Shopper Testing In Home Appraisals Exposes Racial Bias Undermining Black Wealth*. National Community Reinvestment Coalition. <https://nrcr.org/faulty-foundations-mystery-shopper-testing-in-home-appraisals-exposes-racial-bias-undermining-black-wealth/>

¹⁸ Ambrose, B. W., Conklin, J., Coulson, N. E., Diop, M., & Lopez, L. A. (2021). *Do Appraiser and Borrower Race Affect Mortgage Collateral Valuation?* [SSRN Scholarly Paper]. <https://doi.org/10.2139/ssrn.3951587>

iv. While covering a different activity, The PAVE Report and Freddie Mac's research on appraisals can still provide insight into how Agencies might measure discriminatory bias in AVMs.

An earlier report from Freddie Mac's research division concluded that 12.5 and 15.4 percent of homes in majority Black and Latine census tracts were valued less than the contract price compared to only 7.4 percent of white neighborhoods. Additionally, it showed that as the concentration of Black and Latine residents increased, the likelihood that a valuation would fall below the contract price increased.¹⁹ Such outcomes are a foundational aspect of gaps in the attainment of homeownership.

It is worth underscoring how not all appraisers were equally likely to find values below contract prices. The portfolios of almost half of all appraisers did not reveal statistically significant appraisal gaps.²⁰ The fact that unexplained variations were only present among a subset of appraisers points to appraiser bias as a driver of valuation gaps. But the PAVE report's conclusion also suggests that comparing outcomes of AVM models would protect consumers from unfairness.

These findings underscore the value of supervising decision-maker's actions and should inform Agencies of appropriate measures for how model creators should be required to review their AVMs for fairness.

The Agencies should ask model creators to measure fairness for the selection of comps by users of AVMs. Measurements should include the distance of comps from the property and standard deviations in the difference between the valuation and the contract price. As well, our research suggests that appraisals made property markets outside of unincorporated areas are susceptible to capturing the impact of historic redlining. As a result, it would also be important to test AVMs for potential disparate treatment of properties based on the category of incorporation. If there are significant differences, it could be a flag for further review. Such an approach would measure "inputs."

However, measuring outcomes is also necessary. It should consider how values of otherwise similar homes in neighborhoods of color are valued compared to the assessments made by AVMs in majority-white communities.

Disparate impact theory is an appropriate framework to review AVMs for discriminatory impacts.

In ECOA, regulators have applied the theory of disparate impact to identify instances where discriminatory practices may exist in lending. Appraisals and evaluations used to assess the collateral values have a meaningful effect on the outcomes of credit decisions.

v. AVMs, like all algorithmic processes, can achieve dual goals of accuracy and fairness.

The Agencies can hold algorithms accountable to achieve fairness and accuracy without compromising one for the other. Such an expectation is necessary to establish a fifth quality control for nondiscrimination. The Agencies should expect AVMs to produce estimates with high confidence of accuracy (control one) and to avoid having discriminatory impacts (proposed quality control five).

¹⁹ Freddie Mac. (2021). *Racial and Ethnic Valuation Gaps In Home Purchase Appraisals* [Research Insights Notes and Briefs]. <https://www.freddiemac.com/research/insight/20210920-home-appraisals>

²⁰ Freddie Mac. (2021). *Racial and Ethnic Valuation Gaps In Home Purchase Appraisals* [Research Insights Notes and Briefs]. <https://www.freddiemac.com/research/insight/20210920-home-appraisals>

We are optimistic because debiasing techniques can make models fairer and more accurate.²¹ Fairness (the fifth quality control) and accuracy (the first control) are in accord with proper model supervision.²²

vi. Creators of AVMs would benefit from clarity on how they can self-test their models for discrimination.

To support the implementation of the control standard for testing AVMs, the Agencies should take steps to improve self-testing by AVM users.

ECOA permits creditors to request demographic data in mortgage and small business lending contexts. Lenders solicit information for HMDA and Section 1071 small business data reporting rules. A benefit to these activities is that they provide inputs to facilitate underwriting model testing for potential discriminatory impacts.

The same capabilities are not possible for testing AVM models. Lenders cannot share demographic data with AVM providers.

If users of AVMs could have loan-level demographic data, it would facilitate testing programs. However, to accomplish that result, those users will need clarity on how they can ask for demographic information on credit-seekers, be it in the context of a home mortgage, HELOC, loan modification request, or any other listed use case.

The Agencies should indicate that soliciting demographic data for testing is permissible. To facilitate it, they should clarify how requests should be made. Agencies should provide model forms for soliciting information, for example. Additionally, testers should know what are appropriate sample sizes. Sample sizes should be large enough to provide statistically significant conclusions across diverse populations.

As well, the Agencies should provide interested non-governmental groups with data that will create the opportunity to conduct testing on the fairness of AVMs. A precedent exists to show how public access to mortgage lending data can improve markets. Specific unit-level information on the property's profile would be necessary to fulfill the needs of third parties. Second, third parties would require information on the property's census tract to support reviews of potential redlining. Census tracts would have descriptors for demographic data on that neighborhood. Needed data would include all relevant aspects of the property (bedrooms, bathrooms, square footage, lot size, property type, et al.) useful to most AVMs. The data would include demographic descriptors for the property's census tract. Third, it would have demographic data on the household requesting the evaluation.

vii. Agencies should release data to the public to allow third parties to review AVM models for fairness.

The Agencies should provide interested non-governmental groups with data that will create the opportunity to conduct testing on the fairness of AVMs. A precedent exists to show how public access to mortgage lending data can improve markets. Specific unit-level information on the property's profile would be necessary to fulfill the needs of third parties. Second, third parties would require information on the property's census tract to support reviews of potential redlining. Census tracts would have descriptors for demographic data on that neighborhood. Needed data would include all relevant aspects of the property (bedrooms, bathrooms, square footage, lot size, property type, et al.) useful to most AVMs. The

²¹ Singh, A., Singh, J., Khan, A., & Gupta, A. (2022). Developing a Novel Fair-Loan Classifier through a Multi-Sensitive Debiasing Pipeline: DualFair. *Machine Learning and Knowledge Extraction*, 4(1), Art. 1. <https://doi.org/10.3390/make4010011>

²² Wick, M., Panda, Swetasudha, & Tristan, J.-B. (2019). Unlocking Fairness: A Trade-off Revisited. *Advances in Neural Information Processing Systems*, 32. https://papers.nips.cc/paper_files/paper/2019/hash/373e4c5d8edfa8b74fd4b6791d0cf6dc-Abstract.html

data would include demographic descriptors for the property's census tract. Third, it would have demographic data on the household requesting the evaluation. To facilitate accountability across the entire lifecycle of lending for dwellings, a similar data set would exist for the secondary market. This data set would facilitate engagement by the public with FHFA, the GSEs, and other major purchasers of securities.

viii. Tests based on outcomes are likely more valuable when reviewers determine if an AVM is discriminatory.

In some cases, it may not be possible to make a scientifically-defensible judgment on the impartiality of a model based solely on an assessment of the inputs it used.

Consider the example of an AVM that uses pictures posted on the internet to make assessments of property value. An AVM could easily "scrape" photos from multiple sources to assemble a data set covering any home listed for sale or rent, and it could supplement that data set with other pictures with geolocation data tagged in their metadata.

If trained on such a data set, an AVM could learn bias. With minimal model training, a company could create an AVM model to assess home values using inferences based on factors with implicit demographic associations. Even if the model excluded family photos from consideration, it could still utilize elements such as home décor²³ (African wall hangings, wall photos from the civil rights era) or the presence of assistive living technology (shower bars, wide doorways, outdoor entrance ramps) that could have high correlations to protected class status. Given that there are patterns of home décor preference and protected class status – and residential segregation continues to exist – such elements could reaffirm prior appraisal biases.

This example underscores the value of reviewing models based on their outcomes. Returning to the earlier example of a model that used photographs to evaluate home values, it is pertinent to understand that photographs are an example of unstructured data. They do not fit inside an Excel file. They are hard to evaluate for bias on their own. There is no way to derive a score for a photograph or a collection of photographs that will fit inside a regression model. However, the neural networks for which they are likely to be used as inputs can be reviewed for the conclusions they make.

The possibility has ramifications for when an independent testing agency reviews a model. It would be far easier to detect the impacts of such a model on its outcomes. A tester would need as many reference points from independent research as there are variables in a model. Moreover, inputs are blind to model weights, and even more so when a model uses decision trees and neural networks.

The Agencies should use existing testing technologies to ensure AVMs work as intended. AVMs should be reviewed for discriminatory effects at all stages – when they are built, as they are trained, and after they have finished their assessments.

ix. Reviews should ensure that models are trained on representative data sets. Reviews can use statistical procedures to determine which inputs contribute to discrimination.

Data sets used for training should include homes from neighborhoods with diverse populations. If a model does not include a fully-representative cross-section of housing, it could introduce errors in

²³ Lange, A. (2023, March 9). The Interior Lives of Black Homes. *The New York Times*.
<https://www.nytimes.com/2023/03/09/style/black-home-interior-design.html>

valuations. Reviews of models should ensure that they have adequate numbers of homes from neighborhoods composed of diverse communities.

Because models can "drift" over time, testers should review models made by companies that apply machine learning techniques to their AVMs regularly. Machine learning is not inherently bad or good; sometimes, it may be a valuable tool to prevent model drift. Nonetheless, regulators should expect companies to monitor for drift.

Testing for drift has value even if models do not change. Even if a model remains constant, societal factors may evolve that change the accuracy of a model. COVID-19 provides an excellent example of this possibility. The economic disruption brought about by the pandemic destabilized the financial health of otherwise credit-worthy households. In the future, models unchanged since before the pandemic might make incorrect conclusions about valuations.

We also support the Agencies' second quality control, which will insist that AVMs adhere to quality control standards to prevent data manipulation. The Agencies should clarify that a training data set that is not diverse may not meet quality control standards designed to avoid data manipulations. The Agencies should also sample training data sets to ensure that these inputs are consistent with demographic shifts and alterations in preferences for certain types of housing.

x: Appraisals made by an AVM should be explainable.

Regulation B requires creditors to provide copies of appraisals and other written valuations developed in connection with an application for a loan secured by a first lien on a dwelling.²⁴ The CFPB clarified that this expectation applies to credit underwriting models using algorithms in May 2022.²⁵ The Agencies should confirm that results from AVMs should also be explainable.

Currently, all mainstream methods for making appraisals (sales-based, income-based, and cost-based) do so within an easily-explained framework. Appraisals show the address, features, and weights associated with comparable properties.

AVMs may not use a linear approach, may use thousands of variables, and rely on dynamic weightings. As a result, efforts to explain their reasoning must be flexible and nuanced. The system of reason codes used for current adverse-action rules needs to be better suited to explain credit underwriting decisions, and it is similarly inadequate for this purpose.

As a result, model disclosure forms must be flexible for AVM users to "write in" the most determinative factors to their reasoning.

Secondly, the agencies should require lenders to use AVMs in connection with a credit decision to use an explainability technique to ensure consumers understand the basis for a valuation. Explanations of credit decisions can use Shapley values. Shapley explanations can explain "local" outputs. They reveal which

²⁴ Consumer Financial Protection Bureau. (2013, January 31). Disclosure and Delivery Requirements for Copies of Appraisals and Other Written Valuations Under the Equal Credit Opportunity Act (Regulation B). Final Rules. <https://www.consumerfinance.gov/rules-policy/final-rules/disclosure-and-delivery-requirements-copies-appraisals-and-other-written-valuations-under-equal-credit-opportunity-act-regulation-b/>

²⁵ <https://www.facebook.com/CFPB>. (2022, May 26). *CFPB Acts to Protect the Public from Black-Box Credit Models Using Complex Algorithms*. Consumer Financial Protection Bureau. <https://www.consumerfinance.gov/about-us/newsroom/cfpb-acts-to-protect-the-public-from-black-box-credit-models-using-complex-algorithms/>

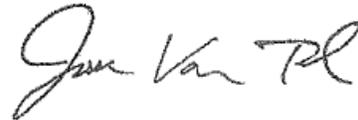
factors made the most significant contribution to a specific outcome.²⁶ Shapley methodology will work equally well for explaining the basis for an appraisal valuation.

Conclusion

Thank you for the opportunity to comment on this emerging issue in housing finance. The Agencies are correct to act to ensure the use of automation valuation models conforms to standards for accuracy, prevents data manipulations and conflicts of interest, monitors for appropriate statistical methods, and corrects any discriminatory elements.

Please contact me or Adam Rust (arust@ncrc.org) with any questions or needs for clarification.

Sincerely,



Jesse Van Tol
Chief Executive Officer
National Community Reinvestment Coalition
jvantol@ncrc.org

²⁶ Laura Blattner, Jann Spiess, & P-R Stark. (2022). *Machine Learning Explainability & Fairness: Insights from Consumer Lending* [Working Paper]. FinRegLab and Stanford University. https://finreglab.org/wp-content/uploads/2022/04/FinRegLab_Stanford_ML-Explainability-and-Fairness_Insights-from-Consumer-Lending-April-2022.pdf

CC:
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/Legal OES
Federal Deposit Insurance Corporation
550 17th Street NW
Washington, DC 20429

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

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

Clinton Jones
General Counsel
Attention: Comments
Federal Housing Finance Agency
Fourth Floor, 400 Seventh Street SW
Washington, DC 20219

Comment Intake
Consumer Financial Protection Bureau
c/o Legal Division Docket Manager
1700 G Street NW
Washington, DC 20552