Fannie Mae®

Appraising the Appraisal

A closer look at divergent appraisal values for Black and white borrowers refinancing their home

Jake Williamson
Mark Palim
February 2022

This paper, originally published in January 2022, has been updated to provide additional clarification on the research methodology.

Findings

Black borrowers refinancing their home on average received a slightly lower appraisal value relative to automated valuation models, according to an analysis of 1.8 million appraisals conducted as part of refinance applications in 2019 and 2020. At the same time, white borrowers refinancing their home on average received a slightly higher appraisal value relative to the models. This was true for homes in both majority-white and majority-Black neighborhoods.

Homes owned by white borrowers were more frequently overvalued than homes owned by Black borrowers. Overvaluations of white-owned homes were present at a higher rate in all neighborhoods, but were more likely to occur among homes owned by white borrowers in majority-Black neighborhoods. Overvalued equates to an appraised value at least 10% higher than the automated valuation models.

The frequency of "undervaluation" did not have a notable racial pattern. Undervaluation of Black-owned homes occurred at the same rate as white-owned homes in majority-white neighborhoods. In majority-Black neighborhoods, the frequency of undervaluation for Black and white homeowners is within a percentage point of each other. "Undervalued" equates to an appraised value at least 10% lower than the automated valuation models.

Six states accounted for nearly 50% of the overvalued homes of white owners in majority-Black neighborhoods. The states were Georgia, Louisiana, South Carolina, North Carolina, Mississippi, and Alabama.

Introduction

Fannie Mae is conducting a multi-part research effort to better understand appraisal valuation differences across different demographic groups.

> While most studies have focused their research at the neighborhood or census tract level, Fannie Mae targeted its appraisal research to isolate valuation differences at the property and borrower level, along with further analyses that focused on the racial composition of neighborhoods.[1]

Specifically, we have examined the extent to which property valuations differ in majority-Black neighborhoods and the extent to which properties owned by Black and white borrowers in these neighborhoods are valued differently, taking into account property characteristics that are normally considered in an appraisal.

For this study, Fannie Mae analyzed 1.8 million appraisals conducted as part of refinance applications from 2019 – 2020, in which homeowners self-identified as Black or white non-Hispanic.[2,3] We chose to study refinance applications, as opposed to home purchase applications, because the appraiser in a refinance transaction typically interacts directly with the homeowner (i.e., the borrower), establishing a pathway for potential bias to influence the appraisal results. The race or ethnicity of the borrower is often disclosed in the loan data, making it possible to directly observe any correlation with value. On the other hand, in a purchase transaction, the appraiser typically does not interact with the buyer (i.e., the borrower) of the property but rather with the seller or the seller's agent. The availability of racial or ethnic data of sellers and real estate agents is limited, thereby making an analysis of valuation differences by different demographics for purchase transactions limited or incomplete relative to the analysis detailed below using refinance transactions. Additionally, Fannie Mae research has shown for purchase transactions when a home appraises below the contract price, this often leads to a renegotiation of the purchase favoring the buyer. Further, this same research concludes that financially

1 When considering the racial composition of an area, we are defining neighborhood as a census tract, consistent with standard economic practice. Census tracts may not necessarily correspond to the appraiser's definition of a "neighborhood."

2 "Black" also includes borrowers who identify as Black in addition to any other race/ethnicity (for example, a borrower who identifies as Black and Hispanic). "White" in this paper refers solely to white non-Hispanic, consistent with industry practice for fair lending analysis. White non-Hispanic will be referred to as "white" for the remainder of this article.

3 When a loan application had multiple borrowers, if any of the borrowers self-identified as a given race (e.g., Black, Asian, Native American, or Hawaiian) or ethnicity (e.g., Hispanic), then the loan application was included in that demographic, consistent with industry practice for fair lending analysis.

constrained borrowers benefit from this renegotiation and are more likely to achieve a renegotiation when an appraisal is valued lower than the contract price.[4] In other words, appraisals below pending contract price, if accurate, do not necessarily equate to an issue.

To evaluate the appraisal results, the values from these appraisals were compared against two automated valuation models (AVMs). These Fannie Mae proprietary AVMs use different modeling frameworks, analytical techniques, and data inputs, providing mutually independent estimates of a property's value.[5] We have utilized them for many years in various valuation oversight functions in both origination and default processes. Specifically, the AVM 1 comparison is utilized to evaluate the potential misvaluation risk of appraisals in the origination process, and follows similar steps in its modeling framework as an appraiser (data assessment/ integrity checks, comparable selection, comparable adjustment, and reconciliation). The AVM 2 comparison is utilized for various portfolio management functions and leverages a machine-learning framework that is fundamentally different than the AVM 1 comparison. Coverage, accuracy, and precision of any AVM varies, and very few AVMs have been tested through multiple financial cycles; however, both of these AVM benchmarks are subject to extensive testing, continuous improvement efforts, and rigorous model governance, which gives us confidence in their reliability for this study and during the applicable time period of the transactions in the dataset.

Additionally, this research is not an attempt to measure value accuracy, per se. Instead, it uses the AVMs as neutral benchmarks to compare appraisal opinions against. Then, the frequency and degree of difference between the AVMs and the appraisal values can be analyzed by the race of the homeowner as well as the racial demographics of the neighborhood where the home is situated.

The results presented here are distributional summaries of the statistical differences in values provided in appraisals versus those from AVMs, and as such are indicative of possible bias but not conclusive. In ongoing work, we are evaluating the impact of additional potential explanatory factors such as gentrification on the reported differences in this study.

Our commitment to racial equity

Fannie Mae is firmly committed to promoting racial equity in housing, including helping all homeowners receive an impartial appraisal. As described in Fannie Mae's September 2021 commentary, Our Commitment to Reducing Appraisal Bias, Fannie Mae is taking a number of steps to combat appraisal bias, including performing research to analyze the prevalence of bias in the appraisal process, and this initial paper is one step of that ongoing research effort.

4 "When the appraised value falls below contract, however, even by a small amount, renegotiation rates rise to over 50%; and rates of renegotiation continue to increase, up to 80%, as appraised values fall further below contract." Hamilton Fout, Nuno Mota, and Eric Rosenblatt, "When Appraisers Go Low, Contracts Go Lower: The Impact of Expert Opinions on Transaction Prices," The Journal of Real Estate Finance and Economics (2021), pp. 1 – 41, DOI: 10.1007/s11146-020-09800-6.

5 Other AVMs and methodologies exist to determine valuations; however, these two were selected because of their unique approaches. Further work could consider the usefulness and sensitivity of the results to weighting AVMs, incorporating other AVM approaches, or even using other methodologies like adjusting with price indices. Also, these AVMs are calibrated to single-family residential housing and different valuation methods are used for other kinds of property types.

A renewed focus on appraisals

In the past year, news media reports alleging racial bias in home appraisals have turned a spotlight on the appraisal process in the mortgage industry. Several published research reports[6] focused on the topic of appraisal bias with varying perspectives on its frequency and severity.

As stated above, we relied on Fannie Mae's proprietary AVMs as neutral benchmarks to evaluate the opinion of value provided by an appraiser on a specific property. These models were built utilizing Fannie Mae's appraisal database that contains more than 54 million appraisals.[7] Fannie Mae's AVMs have been tested and found to be reliable tools for generating property values and in evaluating human appraisals. Both the appraisals and AVMs are relevant and representative of the research target population used to compare the appraisal values.

While Fannie Mae's appraisal database is extensive, the sample used in this study was restricted to data from 2019 to 2020 to focus on more recent transactions. The appraisal data was merged with data from Fannie Mae's Desktop Underwriter® system to obtain the race and ethnicity of the borrower. Then, the appraisal data was filtered to select appraisals for which the borrower's race and ethnicity were identifiable and for properties located in census tracts with majority population indicators, based on the American Community Survey 2018 five-year population estimates. The estimated property values from the AVMs were used to analyze potential undervaluation or overvaluation in appraisals.

White-owned homes received higher appraisals relative to automated valuation models

To evaluate linkages between appraiser valuations and the race of homeowners, we compared the degree of difference between the appraiser opinion and our AVMs, both in the aggregate as well as in majority-Black and majority-white neighborhoods.

6 Andre Perry, Jonathan Rothwell, and David Harshbarger, *The Devaluation of Assets in Black Neighborhoods* (Washington, D.C.: Brookings Institution, 2018), brookings.edu/wp-content/uploads/2018/11/2018.11_Brookings-Metro_Devaluation-Assets-Black-Neighborhoods_final.pdf; Edward Pinto and Tobias Peter, "How Common is Appraiser Racial Bias?" American Enterprise Institute, January 4, 2021, aei.org/how-common-is-appraiser-racial-bias.

7 Digital appraisal data compliant with the industry-standard Uniform Appraisal Dataset.

The histogram in Figure 1 shows the distribution of the differences between the appraisal and the AVM for each combination of homeowner and neighborhood to provide a better sense of the results beyond what can be gleaned from the medians. As can be seen in Figure 1, the median difference in values for the white borrowers in both the majority-Black and majority-white neighborhoods were slightly but consistently positive. That is, homes owned by white borrowers, in the aggregate, tended to have appraisals slightly higher than the AVMs, regardless of the demographics of the neighborhood. At the same time, the median appraisal value for all Black borrowers in the dataset was slightly negative — they tended to be lower than the AVMs. Lastly, and specifically for the white in majority-Black distribution curve, a larger portion of the differences between the appraisal and the AVM were concentrated in the greater than 10% range, indicating a larger volume of appraisals were higher than the AVM benchmark for this demographic versus the other demographic populations. For more statistical information on the distribution, refer to the Appendix.

Figure 1: Distribution of AVM 1 comparison differences by racial groups

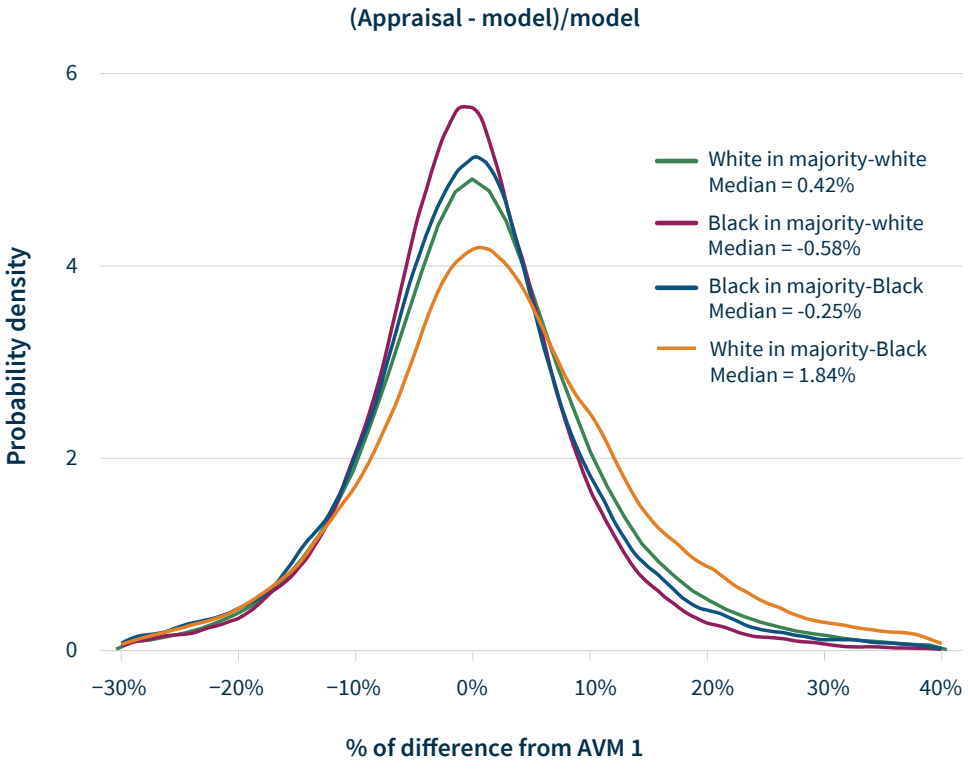


Table 1 provides the dollars and percentages of the median differences between the appraisal and the AVM estimates. Although these figures are relatively small in magnitude, tests measuring the statistical significance of these results confirmed the relevance of the observed differences in value between the appraisal and the AVM. For more information, refer to the Appendix.

Table 1: Median differences from AVM by racial group

Cohort	AVM comparison 1		AVM comparison 2	
	Median: Appraisal — AVM estimate	% AVM estimate	Median: Appraisal — AVM estimate	% AVM estimate
Black in majority-white	($1,693)	-0.58%	($1,143)	-0.39%
Black in majority-Black	($572)	-0.25%	($1,437)	-0.63%
White in majority-white	$1,277	0.42%	$2,026	0.65%
White in majority-Black	$3,955	1.84%	$1,525	0.63%

Where racial disparities in appraised value are most pronounced

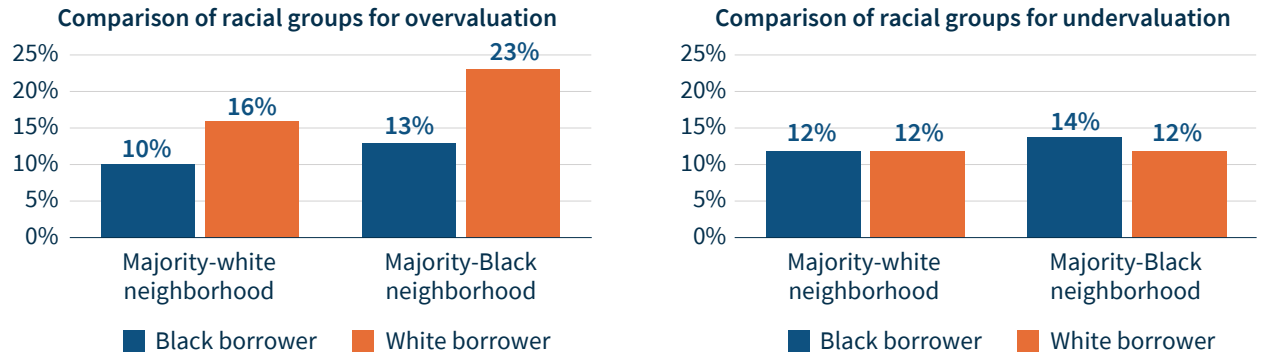
Across the entire dataset, the median differences between appraisal and AVM value are relatively small. For instance, a percentage-point difference between the value of an individual appraisal and an automated valuation model is generally not necessarily considered a red flag for any individual loan.

However, a distinct result emerged when we examined loans at the far ends of our distribution, in particular the population of appraisals that were overvalued. For the purposes of this study, an overvalued appraisal was defined as the appraisal's result being at least 10% greater than the AVM benchmark, while undervaluation was defined as the appraisal's result being 10% or more below the AVM benchmark. We chose 10% as a starting threshold based on our observational experience performing collateral quality control reviews of hundreds of thousands of appraisals since the global financial crisis.

> As shown in Figure 2, the incidence of overvalued appraisals was highest for white-owned homes in majority-Black and majority-white neighborhoods.

Specifically, in majority-Black neighborhoods, white-owned properties were overvalued 10 percentage points more frequently than Black-owned properties. The difference in the frequency in overvaluations was not repeated for undervalued properties. As noted above, this difference in frequency of overvaluations along race may well be due to factors other than racial bias in appraisals, such as gentrification. Testing for such confounding factors is the focus of our ongoing research. The differences observed in undervaluation between white and Black borrowers were similar in rate and not meaningfully different, regardless of the racial demographics of the neighborhood.

Figure 2: Incidence of overvalued and undervalued homes by race of owner



Because AVMs have a margin of error, Figure 2 should not be interpreted as a measure of the frequency of inaccurate appraisals. Some disagreement between appraisals and AVMs should be expected, but the disagreement should be similar across demographic groups. For more information about these comparisons, refer to the Appendix.

The concentration of overvalued white-owned homes in majority-Black neighborhoods also had a distinct geographic element. As shown in Table 2, 49.2% of the overvalued, white-owned properties in majority-Black neighborhoods were concentrated in six states: Georgia, Louisiana, South Carolina, North Carolina, Mississippi, and Alabama. For reference, these six states contain 40.5% of white borrowers in our sample who reside in majority-Black neighborhoods.

Table 2: States with the largest concentration of appraisals overvalued by 10% for white borrowers in majority-Black neighborhoods

Top states		State's share of overall white-owned homes in Black neighborhoods in the sample	
State	**%**	**State**	**%**
GA	14.5%	GA	12.9%
LA	9.0%	LA	6.9%
SC	7.7%	SC	6.5%
NC	6.4%	NC	7.2%
MS	6.0%	MS	3.6%
AL	5.7%	AL	3.4%
Total	**49.2%**	**Total**	**40.5%**

How properties become overvalued

To begin to identify the potential causes of overvalued appraisals, a representative sample of recent refinance appraisals from July to September 2021 loan acquisitions that Collateral Underwriter® (CU®) flagged as overvalued was assessed.

> Collateral Underwriter is Fannie Mae's proprietary automated appraisal risk assessment tool that can evaluate risk on more than 95% of the appraisals submitted for single-family loans acquired by Fannie Mae.

For white borrowers in majority-Black neighborhoods, the leading, separately identifiable reason category of overvaluation is "comparable location." This indicates that the appraiser relied on comparable sales from outside of the subject property's immediate area even though potentially more appropriate comparable properties were available closer to the subject property. By contrast, for Black-owned homes in majority-Black neighborhoods, the leading, separately identifiable reason category of overvaluation was "comparable selection." This means that the comparable sales selected by the appraiser were not the most similar to the subject property in terms of physical characteristics. Table 3 shows the top CU overvaluation reason codes for both white and Black borrowers in majority-Black neighborhoods.

Table 3: Top reasons for overvalued appraisals for white and Black borrowers in majority-Black neighborhoods[8]

White borrowers in Black neighborhoods	
CU overvaluation reason codes	**%**
Multiple factors	36.4%
Comparable location	16.5%
Market adjustments	8.3%
Comparable selection	4.1%
Room count GLA adjustments	5.8%
All other	28.9%
Total	**100.0%**

Black borrowers in Black neighborhoods	
CU overvaluation reason codes	**%**
Multiple factors	29.3%
Comparable selection	12.1%
Market adjustments	9.5%
Comparable location	8.6%
Condition quality adjustments	8.6%
All other	31.9%
Total	**100.0%**

Data: Refinance loan acquisitions from July – September 2021

In addition to using the AVM comparisons and studying the distribution of the differences, Fannie Mae's valuation analysts (who are also certified real estate appraisers) manually reviewed a targeted sample of over 2,000 appraisals from the 2019 – 2020 dataset, including appraisals that were 10% or lower than the AVM 1 comparison, 10% or higher than the AVM 1 comparison, and appraisals within 10% of the AVM 1 comparison. This review provided a quality check for accuracy as a calibration to the AVM 1 comparison results. The internal review results were consistent with the AVM selection of comparable properties and the observations noted for misvaluation.

Our manual reviews observed that, for appraisals that were flagged as having overvaluation risk, the appraiser tended to incorrectly identify neighborhood boundaries, which led to the appraisals having errors in the property value trend (stable, increasing, or declining). For undervalued homes, the appraisals tended to have both inaccurate functional and external property characteristic descriptions, leading to values that were understated. In both cases of overvaluation and undervaluation, the appraisals had similar issues in comparable selection (too distant or dissimilar) and comparable adjustments for specific differences in condition/quality, age, off-street parking, and amenities.

Through the combination of technology and human processes, Fannie Mae and our lender partners prudently manage valuation risk. Our in-house Fannie Mae appraisers perform tens of thousands

8 Multiple factors refer to appraisals flagged for overvaluation but did not identify a single specific reason code.

of valuation reviews each year, which continuously builds our knowledge base to inform our policies and appraiser engagement, as well as to calibrate our AVMs. As part of this work, our engagement strategy includes a dedicated Appraisers web page with our Noble Appraiser video training series and other resources. Several of these resources address comparable selection and comparable location best practices.

Conclusion

Racial inequities in housing are long-standing. Though outright racial discrimination in housing has been outlawed for more than 50 years, its legacy effects continue to be felt in many majority-Black communities and by Black families across several generations. Ensuring that appraisals are as accurate and as unbiased as possible is just one way the housing finance industry can remedy this legacy.

There are many possible reasons why an appraisal would be either overvalued or undervalued. Fannie Mae and its industry partners should continue to take concrete actions to minimize the chances that the race of a borrower or homeowner is one of those reasons. Among these actions:

- **Increase the use of alternative-scope property valuation approaches.** Desktop appraisals develop valuations using a computer with access to a myriad of data and resources — without a physical inspection of the property by the appraiser. So-called hybrid appraisals are similar to desktop appraisals, but require a physical inspection performed by a vetted third party, such as a real estate agent, inspector, or appraiser trainee. Both approaches have the benefit of reducing contact between borrowers and appraisers, thus reducing the likelihood of valuations being affected by personal or unconscious biases.

- **Build on existing safeguards to detect valuation errors.** Automated underwriting platforms should be tightly integrated with systems that can detect asset valuation issues. For instance, Fannie Mae's Collateral Underwriter provides lenders tools that promote greater appraisal quality and help detect a potential misvaluation in real time.

- **Continue to modernize the appraisal process for home loans.** Appraisals are a vital part of mortgage lending; however, continuous improvement through better use of data, technology, and process design can be a powerful way to drive more factual, objective, accurate, and reproducible appraisals.

- **Foster diversity in the appraiser workforce.** Efforts currently underway to expand the participation of people of color in the appraiser profession, including the Appraiser Diversity Initiative sponsored by the Appraisal Institute, Fannie Mae, Freddie Mac, and the National Urban League, should be expanded.

- **Enhance the tools appraisers use to validate their opinions.** Given that many misvaluations are related to items like comparable selection, comparable location, and adjustments, it's important that individual appraisers have the same tools necessary to validate these measures as institutions such as Fannie Mae. Given Fannie Mae's breadth of data and analytical capabilities, providing tools, resources, and information to appraisers to help them more accurately select comparable properties and conduct adjustments will only strengthen and provide more confidence in the appraisal process and product.

About the authors

Jake Williamson is Fannie Mae's Senior Vice President of Collateral Risk Management and oversees its Single-Family Business collateral risk management, loan quality, and operational risk teams.

Mark Palim is Fannie Mae's Vice President and Deputy Chief Economist.

See Appendix on the following page.

Appendix

Additional information for Figure 1

Summary statistics (appraisal % of difference from AVM 1)

	Black in majority-white	Black in majority-Black	White in majority-white	White in majority-Black
Mean	-0.0053	-0.0015	0.0098	0.0290
Median	-0.0058	-0.0025	0.0042	0.0184
Standard deviation	0.0929	0.1090	0.1108	0.1354
Skewness	0.1761	0.2970	1.3395	1.0456
Kurtosis	7.7316	6.9405	52.6354	10.4189
N	55,479	34,193	1,653,792	20,548

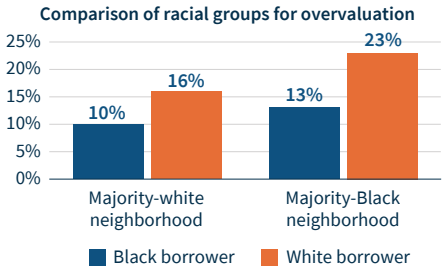
Additional information for Table 1

	Median differences from AVM by racial group					
	AVM comparison 1			AVM comparison 2		
	N	appraisal -estimate	%	N	appraisal -estimate	%
Black in majority-white	55,479	(1,693)	-0.58%	55,479	(1,143)	-0.39%
Black in majority-Black	34,193	(572)	-0.25%	34,193	(1,437)	-0.63%
White in majority-white	1,653,792	1,277	0.42%	1,653,792	2,026	0.65%
White in majority-Black	20,548	3,955	1.84%	20,548	1,525	0.63%

H0: Black in majority-white vs. white in majority-white	p-value = 0.000	**H0: Black in majority-white vs. white in majority-white**	p-value = 0.000
H0: Black in majority-Black vs. white in majority-Black	p-value = 0.000	**H0: Black in majority-Black vs. white in majority-Black**	p-value = 0.000

Additional information for Figure 2

	Comparison of racial groups for overvaluation			
	Majority-white neighborhood		Majority-Black neighborhood	
	Black borrower	White borrower	Black borrower	White borrower
Mean	10%	16%	13%	23%
N	55,479	1,653,792	34,193	20,548
p-value	0.000		0.000	



	Comparison of racial groups for undervaluation			
	Majority-white neighborhood		Majority-Black neighborhood	
	Black borrower	White borrower	Black borrower	White borrower
Mean	12%	12%	14%	12%
N	55,479	1,653,792	34,193	20,548
p-value	0.0521		0.0001