



Reducing Racial Bias in Home Appraisals Using Automated Valuation Technology

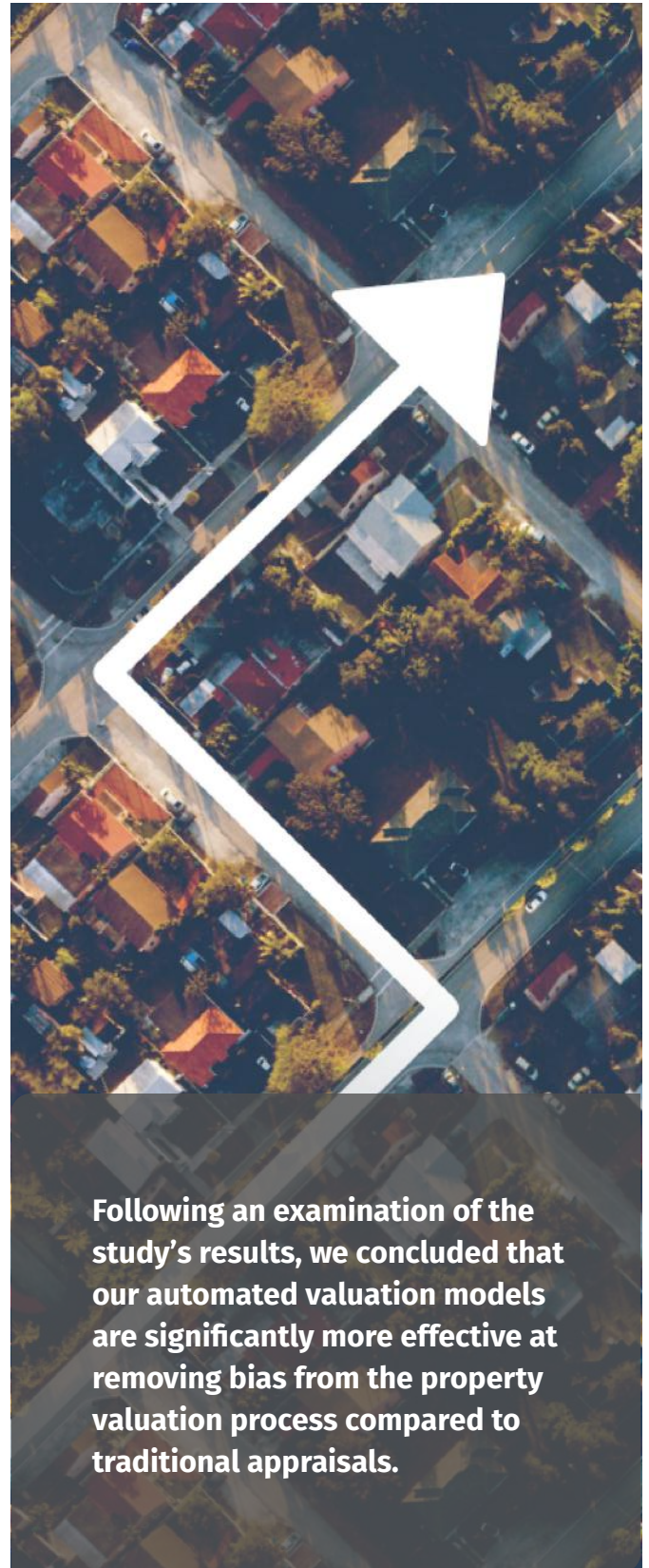


Overview

In September 2021, government-sponsored enterprise Freddie Mac published the results of a study, titled "[Racial and Ethnic Valuation Gaps in Home Purchase Appraisals](#)," which found severe bias present in appraisal data, concluding that – through the traditional appraisal process – racial minorities are more likely to receive an appraisal value that is lower than the contract price the property was sold for.

Traditional appraisals have come under re in recent years for consistently providing inaccurate or insufficient property information and specifically, undervaluing minority-owned homes relative to white-owned homes. When evaluating each property, appraisers must select comparable properties in adherence with a set of rules (e.g., time of sale, distance from target property, etc.) upon which to base their valuation gure. In its Q3 2021 examination of the appraisal industry, Fannie Mae reported that the three most frequent issues in traditional appraisals – which accounted for 62% of the overall appraisal defects – are all related to the selection of comparable properties ("comp selection"). Comp selection is one of the most subjective elements of the traditional appraisal process and is, therefore, highly susceptible to the conscious and unconscious biases of individual appraisers.

HouseCanary has built and developed industry-leading property valuation technology that provides highly accurate, objective information to power real estate transactions of all kinds. We believe our automated valuation models (AVMs), which deliver results based on hundreds of property data points and deep contextual information, can help significantly reduce the bias present in traditional home appraisals. To measure the accuracy of our automated valuation tools in appraising homes in minority neighborhoods, we conducted a statistical study of our technology using the Freddie Mac report as a guide. In particular, we sought to determine whether use of our automated property valuation tools could mitigate the effects of appraisal bias highlighted in the Freddie Mac analysis.



Following an examination of the study's results, we concluded that our automated valuation models are significantly more effective at removing bias from the property valuation process compared to traditional appraisals.

Key takeaways & results

AVMs are more reliable tools for providing a fair valuation when compared to traditional appraisals.

Our results, which can be found in the appendix, show that HouseCanary's AVM tool generally outperformed our comp valuation tool in producing unbiased valuations. When sufficient data is available to support an AVM for a property, our research suggests it is a more reliable tool for providing a fair valuation than either our comp valuation tool or a traditional appraisal.

No evidence of racial bias exists in HouseCanary's automated comp and AVM tools. Following our analysis, which can be found in the appendix, we found no evidence of racial bias in HouseCanary's automated comp and AVM tools. This stands in stark contrast to the results of Freddie Mac's examination of traditional appraisals, which found that "Black and Latino applicants receive lower appraisal values than the contract price more often than White applicants."

HouseCanary's comp-based and AVM values are fair to homeowners when neighborhoods are more diverse. Our research found that our comp-based and AVM values tend to become more favorable to homeowners as the percentage of racial minorities in a tract increases. On the contrary, Freddie Mac found that in traditional appraisals, the discrepancy between appraisal value and sales price "increases as the percentage of Black or Latino people in the tract increases." For example, Freddie Mac's research demonstrated how, as the threshold of minorities increases (e.g., from 50% to 80%), traditional appraisal data becomes more unfair, whereas HouseCanary's models actually tend to favor minority homeowners more as the percentage of minority homeownership in a tract increases.



Conclusion

Having persisted for years, appraisal bias has become an increasingly important issue as countless reports of racial prejudice and discrimination have come to the forefront of public awareness. Earlier this year, the Biden-Harris administration [announced](#) a series of actionable steps to narrow the racial wealth gap in the U.S. and address racial discrimination in the housing market, including inequity in home appraisals.

Amid increased scrutiny on inherent bias in traditional appraisals, some have highlighted the industry's lack of diversity and the steady decline in licensed appraisers, as well as the meaningful progress made in the development of automated valuation tools in recent years.

In 2020, the [Urban Institute](#) found that 89% of all property appraisers and assessors are white while just 2% are Black and 5% are Hispanic. The majority of appraisers fall on the higher end of the age spectrum as well, with very few under the age of 34. The [Appraisal Foundation 2021 Diversity Survey](#) calculated that roughly 66% of all licensed appraisers are above the age of 55. The survey also found that nearly 64% of all appraisers are male.

According to Fannie Mae's estimates, appraisal submission volume has increased steadily over the last decade, reaching record highs in the spring of 2021, while the number of active appraisers has gradually declined at a rate of approximately 1% per year. If this trend continues, automated valuation tools could become all the more important in handling the outsized demand for accurate home appraisals.

The results of our study on automated valuation technology clearly demonstrate that HouseCanary's automated comp and AVM tools are unbiased with regard to race and ethnicity, and can thus augment the appraisal process by curtailing discrimination against minority homeowners.

When used effectively, these tools can also be more efficient and cost-friendly without the implicit bias that appears to be inherent within the traditional appraisal process. With the publication of this research, we look forward to continuing the discussion about the role automated valuation technology can play in the real estate industry to promote equality throughout the entire homeownership journey.



Appendix

Methodology

Data Collection: We began by compiling a random sample of 10,000 arm's length sales of residential real estate that have closed since September 1, 2021. We then removed transactions for which homeowner ethnicity data was unavailable or where we detected issues with the accuracy of integrity of the data, leaving us with a sample size of 9,916 total transactions.

Ethnicity Attributes: To measure the impact of valuation bias across different ethnic groups, we leveraged similar techniques to Freddie Mac. Using census tract level data from the US Census Bureau's American Community Survey, we grouped each tract according to whether the homeowners accounted for were more than 50% White Non-Hispanic, 80% Black or Latino, and more than 80% Minority¹. Following Freddie Mac's approach, if a tract was over 50% White Non-Hispanic, 80% Black, Latino or Minority, our testing assumed that every homeowner in that tract was the same race of ethnicity as the tract's majority group.

One slight difference in our approach versus Freddie Mac is that Freddie analyzed tracts that were 50+% Black, Latino and Minority as well as 80+%. We found that there was a much greater variation of outcomes in the 50+% groups than the 80+% groups, particularly in tracts that were 50+% Minority. In our opinion, the assumption that all homeowners in a tract that is majority non-White can be treated identically is harder to justify at the 50% threshold. So, focusing on 80+% minority tracts seemed likely to yield the most reliable results. In addition, because the Freddie Mac study showed appraisal bias increasing as the percentage of Minority, Black or Latino homeownership increases in a tract, focusing on the tracts that are most vulnerable to appraisal bias in this sense seemed appropriate.

HouseCanary Valuation Tools: Our proprietary valuation tools include an automated valuation model, or AVM, as well as comp valuation tool. Our comp valuation tool begins by examining recently sold properties within a small search radius around the property of interest and slowly expanding the search until 500 appropriate comps have been identified. The tool then ranks these comps based on our proprietary similarity score. Our similarity score is based on inputs such as GLA and bed, bath and lot size as well as the outputs from a regression model that accounts for time of sale as well as distance between properties. Our comp valuation tool calculates a value for the subject property based on the average sale price among the 10 comps with the highest similarity scores.

Measuring Valuation Gap: To test our valuation tools for bias, we began by measuring the percentage difference between the value calculated by our tool for each subject property and the sale price of that property. We did this by subtracting the sale price from the tool value and then dividing the result by the sale price.

$$\text{percent difference} = \frac{\text{Comp Value} - \text{Sale Price}}{\text{Sale Price}}$$

¹Every group other than non-Hispanic Whites (e.g., Asian) was counted as Minority.

Methodology cont.

For our AVM, the equation is:

$$\textit{percent difference} = \frac{\textit{AVM value} - \textit{Sale Price}}{\textit{Sale Price}}$$

We believe sale price is the fairest metric to use as a target variable since this value is agreed upon by two separate parties negotiating at arm's length. In each case, a positive percent change can be interpreted as an overvaluation and a negative value can be interpreted as an undervaluation.

By aggregating these calculations by census tract and comparing the results for different tracts, we can examine whether a relationship appears to exist between the racial or ethnic makeup of a tract and the likelihood of a valuation gap occurring, i.e., whether valuation bias seems to exist.

Results

The results of multiple tests are provided in the following section. We ran a two-tailed t-test to compare mean valuation gaps and a one-tailed t-test where the two-tailed t-test failed.

In the tables that follow, the results of the t-test are provided as well as the gap between the White and Minority tracts. The null and alternative hypotheses are also provided for each of the t-tests. Each t-test and results are reported using a 0.05 significance level. This is used consistently when rejecting or failing to reject a null hypothesis.

These tests start with two baselines: the valuation gap for the total sample and for White tracts. You will find these results in the first two rows of the result tables. The valuation gap is then presented for tracts that are 80+% Minority, Latino or Black. In addition, we ran t-tests comparing the mean gaps of the Minority, Black and Latino tract to the White tracts.

Null Hypothesis (2-Tail): Group mean for valuation gap of Black, Latino or Minority tracts is equal to the group mean for valuation gap of White tracts

Alternative Hypothesis (2-Tail): Group mean for valuation gap of Black, Latino or Minority tracts is not equal to the group mean for valuation gap of White tracts

If any two-tail test was rejected, we ran a one-tail test. The 1-tail hypotheses were:

Null Hypothesis (1-Tail): Group mean for valuation gap of White tracts is less than or equal to the group mean for valuation gap of Minority tracts

Alternative Hypothesis (1-Tail): Group mean for valuation gap of White tracts is greater than the group mean for valuation gap of Minority tracts

Comp Based Valuation Gap Results

Given all tests and p-values in the table, we failed to reject the null hypothesis at the 0.05 significance level for all groups. This shows that there is no evidence of racial or ethnic bias when using the automated comp based valuation

Property Tract	Count	Mean Valuation Gap vs Contract Price	Gap vs. White	P-value
Overall	9916	0.67%	-	-
White	7787	1.2%	-	-
Minority (80%-100%)	633	-0.33%	-1.53%	.3264
Latino (80%-100%)	71	7.59%	+6.3%	.5781
Black (80%-100%)	132	0.30%	-0.9%	.5377

Property Tract	Count	Mean Valuation Gap vs Contract Price	Gap vs. White	P-value
Overall	9916	2.01%	-	-
White	7787	2.31%	-	-
Minority (80%-100%)	633	2.02%	-0.29%	.8621
Latino (80%-100%)	71	11.2%	+8.89%	.4746
Black (80%-100%)	132	1.92%	-0.39%	.8479

AVM Valuation Gap Results

Similar results are found when measuring bias with our AVM valuations. Given all tests and p-values in the table, we fail to reject the null hypothesis at the 0.05 significance level for all groups. Whether using HouseCanary's AVM or Comp Based Value to generate valuations, there is no evidence of bias against the different minority populations.

Additional Testing

To further demonstrate the robustness of our findings, we performed these same tests using data from our internal AVM prelist benchmark. We use this benchmark to examine our AVM performance by comparing the AVM's pre-listing estimated value for a property that has been sold with its closed sale price. The dataset for these tests consists of all homes that have sold in the previous 6 months (dating back to May 1st 2021). An advantage of testing this benchmark is that it provides a much larger sample size than is possible with the randomized sample based methods used above.

Prelist AVM Valuation Gap

For the Latino 80+% group, we fail to reject the null hypothesis that the mean valuation gap for White tracts and Latino tracts are equal. For Black and Minority tracts, these tests show a positive valuation gap compared to a slightly negative valuation gap for White tracts. So, Black and Minority homeowners actually fare slightly better in these tests than White homeowners. This result reinforces our conclusion that our AVM does not generate values that are biased against Black, Latino or Minority homeowners.

Property Tract	Count	Mean Valuation Gap vs Contract Price	Gap vs. White	P-value
Overall	645,899	-0.8%	-	-
White	503,508	-0.9%	-	-
Minority (80%-100%)	37,899	0.1%	+1.0%	<0.0001 (1-Tail)
Latino (80%-100%)	4,701	-0.6%	+0.3%	0.224 (2-Tail)
Black (80%-100%)	6,419	2.3%	+3.2%	<0.0001 (1-Tail)

References

1. [Freddie Mac Research Note](#)
2. Fannie Mae Q3 Appraisal Update
3. [White House Task Force](#)
4. [Urban Institute](#)
5. [Appraisal Foundation 2021 Diversity Survey](#)