

# How Common Is Appraiser Racial Bias?

An Analysis Using Big Data to Determine Whether It Is Common or Uncommon that an Appraiser's Knowledge of an Applicant's Race Results in Valuation Bias

Edward Pinto ([pintoedward1@gmail.com](mailto:pintoedward1@gmail.com)), Director  
Tobias Peter ([Tobias.Peter@AEI.org](mailto:Tobias.Peter@AEI.org)), Director of Research

AEI Housing Center

Working Report – Preliminary Results

[AEI.org/housing](https://www.aei.org/housing)

January 4, 2021

Link to AEI HMIs:

<https://www.aei.org/housing/housing-market-indicators/>

We grant permission to reuse this presentation, as long as you cite as the source:

AEI Housing Center, [www.aei.org/housing](https://www.aei.org/housing).

# Recent Reports Have Alleged Racial Bias by Appraisers on Mortgage Loans

- New York Times examples\*
  - \$135,000 or 29% difference (Jacksonville FL - appraisal 1: \$330,000, appraisal 2: \$465,000)
  - \$40,000 (% unknown) difference (Hartford, CT suburb)
  - \$160,000 or 24% difference (Los Angeles - appraisal 1: \$500,000, appraisal 2: \$660,000)
- Denver News Channel 7 example\*\*
  - \$145,000 or 26% (Denver, CO – appraisal 1: \$405,000, appraisal 2: \$550,000)
- Chicago Sun Times example\*\*\*
  - \$62,000 or (Chicago, IL – appraisal 1: \$278,000, appraisal 2: \$340,000)
- For the four examples with sufficient data, appraisal 1 came in an average of about \$126,000 or 25% lower than appraisal 2.
- These claims of bias were all based on an allegation that a human appraiser was aware of the applicant(s)' race either from a meeting or photos or other items in the home which indicated race, and as a result, the appraiser underestimated the property's value.
- These allegations of undervaluation commonly occurred in predominantly White neighborhoods.
- The implications of these three stories are that intentional and perhaps unintentional appraisal bias is commonplace and the valuation gaps are large.
- A literature search found no statistical analysis of this type of claim.
- We undertake the first statistical analysis to evaluate whether this type of claim of intentional racial bias, along with unintentional bias, are common or uncommon.

\* <https://www.nytimes.com/2020/08/25/realestate/blacks-minorities-appraisals-discrimination.html>

\*\* <https://www.thedenverchannel.com/news/local-news/an-unconscious-bias-biracial-denver-couple-says-they-faced-discrimination-on-home-appraisal>

\*\*\* <https://chicago.suntimes.com/2020/10/7/21493755/chicago-home-appraisal-black-race-homeowners>

# We Asked the Collateral Risk Network (CRN) to Survey Appraisal Management Companies and Lenders

- Why: our data consist of appraisals on transactions that actually closed, thus we need to evaluate whether our set suffers from selection bias as discrimination might occur with respect to appraisals not used in loan closings.
- The question is: how likely is that Review of Valuation (ROV) requests are received and what happens after an ROV is made? Here are relevant results:

		AMC	Lenders
Frequency of Review of Valuation (ROV) requests:		Insufficient data (only 1 response, as AMCs generally do not see ROV requests)	1st survey- 2.5% (median for 12 responses) 2nd survey- 3% (median for 8 responses)
Most common reasons for a ROV (where one is given):		1. Poor selection of comps 2. Recent Improvements not noted 3. Sq. Ft. incorrect	1. Poor selection of comps 2. Sq. Ft. incorrect 3. Recent Improvements not noted
Regarding ROVs involving racial or ethnic bias:	Never	58%	63%
	<10%	42%	37%
When you do get an ROV, how many get escalated to a second appraisal?	Never	0 (0%)	2 (25%)
	<10%	17 (89%)	4 (50%)
	10-30%	2 (11%)	2 (25%)
	>30%	0 (0%)	0 (0%)
How many contracts get renegotiated?	Never	3 (17%)	0 (%)
	<10%	9 (50%)	5 (71%)
	10-30%	5 (28%)	1 (14%)
	>30-60%	0 (0%)	1 (14%)
	>60%	1 (6%)	0 (0%)
Are ROVs more common on refinance or purchases?		68% refi	63% refi

- Conclusions:
  - Items noted in red are not unique to claims of racial or ethnic bias, but are common to all ROVs.
  - An ROV is rare (2.5%-3% incidence), ROVs are related to race or ethnic bias are also rare, and an ROV is a condition precedent to a second appraisal, which itself is relatively rare (about 10%-20% incidence).
  - Based on the survey, we may conclude that our data set, consisting entirely of closed loans, does not suffer from a significant level of selection bias.

# Home Mortgage Disclosure Act (HMDA) Data Confirms This Conclusion

- HMDA, a virtual census of institutional loan app data, covered >8 million applicants in 2018-2019.

		White	Black
Refi	Total Applications	4,995,603	614,394
	% of Application Denied	16%	26%
	% Denial Reason is Collateral	17%	14%
	% Denial Reason is DTI or Credit History	49%	54%
Purchase	Total Applications	6,065,075	782,255
	% of Application Denied	8%	17%
	% Denial Reason is Collateral	13%	8%
	% Denial Reason is DTI or Credit History	56%	66%

Note: Data are for first lien, 1-unit, owner-occupied homes only. They exclude I/O, balloons, or neg. am. loans.  
 "Denial "is for the principal denial reason only.

Source: HMDA 2018 & 2019

- **Conclusions:**
  - Loan denial rate on refis is 16% for White applicants and 26% for Black applicants; denial rate on purchase loans is lower for both White and Black applicants.
  - The main reasons for Black denials are credit history and DTI.
    - Collateral is noted as a denial reason less for Black applicants than White ones and perceived or alleged bias would be one of a number of potential reasons for a collateral denial.
  - The potential for bias by an appraiser is small, since it would be a percentage of a percentage.
  - Since HMDA is consistent with the results of the CRN survey, we may conclude that our data set, consisting entirely of closed loans, does not suffer from a significant level of selection bias.



# We Use Big Data to Evaluate the Likelihood of Racial Bias by an Appraiser

- The AEI Housing Center has assembled the National Housing Market Database (NHMDB), which uses and connects many different datasets.
  - Most in-depth resource for key housing data and trends.
  - Accurate, timely, and in-depth coverage of purchase trends.
  - Connects the dots for many housing indicators, yielding the most comprehensive analysis of the housing market.
  - Its integrated structure is uniquely suited to study countless questions.
  - Here we apply the NHMDB to help shed light on the question: Is there intentional or unintentional racial bias by appraisers?
    - We are able to identify whether the transaction had a human appraiser or a GSE appraisal waiver using a computer generated value.
      - The waiver value is unbiased, as neither race nor any other borrower characteristic is used that could be correlated with race is used in generating the value.
    - We use the race neutral Automated Valuation Model (AVM), where neither race nor any other borrower characteristic is used that could be correlated with race is used in generating the value.
    - The AVM controls for non-racial factors that affect home value, leaving the race dummies to pick up any racial effects.
    - AVMs allow us to statistically test bias assertions by objectively assessing previously performed human appraisals and appraisal waivers.
  - A literature search found no statistical analysis of this type.
    - This study is based on 243,000 appraisals (of which 59,000 are GSE appraisal waiver cases & 22,000 GSE human appraisals), & 6,500 pair match cases originated in 2018 & 2019.
  - Unintentional bias may have a disparate impact, but it must be demonstrated that a challenged practice involving racial bias is the substantial cause of the disparate impact.
  - Using computer generated property values, we used big data to look for bias of the type alleged in the media and racial bias that might result in a disparate impact allegation.
- This Working Report presents preliminary results and we welcome comments.

# Using Big Data to Evaluate the Presence of Appraiser Bias

- NHMDB data used in performing this analysis include:
  - Public records
  - Automated Valuation Models (AVM)
  - National Mortgage Risk Index using agency loan data, including use of appraisal waivers
  - Home Mortgage Disclosure Act (HMDA) data
- Our main focus is on refinance loans:
  - Refinance loans lack an arm's length transaction, potentially providing more of an opportunity to exercise racial bias.
  - Purchase loan appraisals are highly anchored to the sales price and therefore tend to not come in below the sales price.\* This makes them useful as a check, since they provide less of an opportunity and potential to exercise racial bias.
  - Thus, deviations due to bias, if present, should be larger for unanchored refi valuations compared to anchored purchase valuations.
  - FHA purchase loans have a unique double anchor, which is useful for a natural experiment:
    - The 1<sup>st</sup> anchor is to sales price and the 2<sup>nd</sup> to a statutory maximum LTV of 96.5%.
    - Also, an average LTV of around 95% for all 4 groups being compared, which means even a small gap between the sales price and appraised value would require a price renegotiation with the seller.

\* [Contract Price Confirmation Bias: Evidence from Repeat Appraisals](#)

# Methodology

We use 2 different approaches, to see if Blacks received different results for refinance loan appraisals compared to Whites, especially as it relates to alleged practices.

- Automated Valuation Model (AVM) approach (N-count = 243,000 loans)
  - We regress the Dec-2017 AVM, month of origination dummies, census tract fixed effects, and a dummy variable for minority status on the log property value.
    - An AVM provides a neutral control for neighborhood and home characteristics.
    - Lacks “knowledge” of applicant’s race.
    - We filter out properties with extreme outliers in home characteristics.
  - Our analysis is enhanced by running it separately for loans using human generated appraisals vs. a waiver with a computer generated value (the AVM noted above is independent of this waiver value).
    - The goal is to determine whether an alleged practice (such as racial bias based on knowing whether an applicant is Black from a meeting or interior items (ie. photos, art, or books) occurs more frequently on Black refi valuations by humans.
    - A second goal is to test for unintentional racial bias against Blacks on refi valuations by humans that has a disparate impact and is the substantial cause of the disparate impact.
  - Hedonic approach is used to check the AVM approach:
    - We replace the AVM with house characteristics (sq. footage, lot size, and dummies for year built).
- Pair approach (N-count totaled 2,600 pairs):
  - Identify pairs of homes that are identical based on external characteristics (same census tract, home type, sq footage, lot size, year built, # of baths, and land use code), within 1,650 ft of each other (median: 200 ft), and owned by households of different racial backgrounds.
    - Rules out challenged practices such as comp selection and neighborhood effects as an explanation.
  - We regress a set of dummies for minority status and month of origination on the log property value. The pair of homes are treated as fixed effects.
- We define the following non-overlapping groups (N-counts):

	AVM/Hedonic	GSE Human	GSE Waiver	Pair
White (defined as non-Hispanic White)	212,935	53,397	21,539	3,563
Black (defined as non-Hispanic Black)	30,435	5,393	1,391	2,810

Note: 2% of Hispanics identify as Black Hispanic. These loans are excluded from this analysis.

# AVM Approach Results for Refinance Loans

- The media reports all had a common challenged practice: the human appraiser knew the applicant was Black by having met the applicant or from photos in the home and rendered a biased opinion of value that was below the property’s correct value.
- The race neutral AVM approach was applied to valuations on: 1. refinance loans generally, 2. purchase loans, and 3. refinance loans with either a human appraisal or a waiver.
  - #1 (refinance loans only) includes instances where biased refi valuations have been rendered.
  - #2 (purchase loans only) would be much less influenced by bias due to anchoring to the sale price by human appraisers.
  - #3. a. (waiver only) has lack of “knowledge” of race and interior condition, or the potential for bias in the selection of comps and should be free of racial bias due to the common practice noted above.
- Our results indicate that for #2 (purchase loans) and #3.a. (non-human valuations), Blacks had nearly identical valuation gaps as Whites (-0.8% and -0.5% respectively), as #1 (-0.7%).
- Thus we conclude allegation that knowing the race of the applicant results in racial bias by appraisers on refinance loans is uncommon and not systemic. This same analysis supports the conclusion that unintentional bias based on race is also uncommon and not systemic.

AVM Approach	Gap between Blacks and Whites
1. For refinance loans only	-0.7%***
2. For purchase loans only	-0.8%***
3. Limited to refinance loans with an Appraisal/Waiver Flag (GSE only)	-0.7%***
a. Waiver only	-0.5%
b. Appraisal only	-0.8%***

Note: The results are differences in property value relative to Whites. \*\*\* denotes significance at the 1% level. Data for this approach are limited to property value \$100,000-\$1,000,000, Year built between 2000-2019, Lot sizes of 2,000-20,000 sq. ft., and Building area of 800-3600 sq. ft.

# AVM Approach: Robustness Checks

	Gap btw. Blacks and Whites
AVM Approach for refinance loans	-0.7%***
Check 1. AVM Approach for refinance loans but using quantile regression (median value)	-0.7%
Check 2. AVM Approach for refinance loans but limiting to properties with values between the 20 <sup>th</sup> and 80 <sup>th</sup> percentile at the census tract level	-0.4%***
Check 3. AVM Approach for refinance loans but adding controls for FICO bucket, income, CLTV bucket, and the # of borrowers	0.4%**
Check 4. Hedonic approach for refinance loans	-2.3%***
Check 5. Hedonic approach for purchase loans	-2.4%***

Note: The results are differences in property value relative to Whites. **\*\* denotes significance at the 5% level and \*\*\* at the 1% level.** Data for these approaches are limited to property value \$100,000-\$1,000,000, Year built between 2000-2019, Lot sizes of 2,000-20,000 sq. ft., and Building area of 800-3,600 sq. ft.

- Checks 1 & 2 focus on the impact of outliers, as the anecdotal value gaps averaged 25% and indicate that the AVM Approach results are very robust (that is, remain valid under different assumptions, parameters and initial conditions).
- Check 3, which controls for certain borrower characteristics, eliminates the small gap between Blacks to Whites (now a small positive gap).
  - Generally, higher income, higher FICO, lower CLTV, and 2 borrowers mean a slightly higher appraisal.
  - The results are also generally robust for different states, but n-counts are often too small for a systematic analysis.
- Checks 1-3: the gaps for refinance loans likely reflect unobserved property or location characteristics that are not captured in the AVM.
- Checks 4 and 5 yields a Black-White gap of -2.3% on refinance and -2.4% on purchase loans. Racial discrimination is unlikely on purchase loans due to sale price anchoring.
- Checks 1-5 support our conclusion that the common challenged practice of knowing the race of the applicant results in racial bias by appraisers on refinance loans is uncommon and not systemic. They also support the conclusion that unintentional bias based on race is also uncommon and not systemic.

# Pair Approach Results

Pair Approach	Gap between Blacks and Whites
1. For refinance loans only	-0.4%
2. For purchase loans only	0.0%

Note: The results are differences in property value relative to Whites. **The results are not statistically significant.** Data for this Approach are limited to homes built between 1950 and 2019 and 1-unit SF, condos, townhomes, or PUDs within 500m of each other (median is 200 ft).

- The Pair approach has results that are quite close to the AVM approach for refinance loans and there is only a small difference to purchase loans, for which biased appraisals are unlikely. This further confirms our earlier conclusion that intentional and unintentional racial bias by appraisers on refinance loans is uncommon and not systemic.
- This approach looks at two or more identical homes, so as to reduce the possibility for variation between the pairs.
  - This reduces the likelihood that variation in selected comparable sales, differences in racial make-up of comparable locations, and location within a neighborhood affect the result.
- The n-sizes for this analysis are much smaller, which limits our ability to further “slice” the data.
  - We ended up with around 2,600 home pairs with owners of a different race.
  - The results seem to hold for townhomes or condos only, but n-counts are fairly small.
  - There is no statistically detectable difference between Waivers and Appraisals for identical homes, but n-counts again are fairly small.

# Refi Waiver-Appraisal Gap between Whites (Control Group) and Blacks

The Gap of a Human Appraisal Relative to a Waiver	Using AVM Approach to Measure Gap	N-Count (Appraisal)	N-Count (Waiver)
White (control group as presumably unbiased on race)	-0.8%***	45,155	19,987
Black	-1.4%	4,422	1,271

Note: Borrower controls are FICO and LTV buckets, the number of borrowers, and income. \*\*\* denotes significance at the 1% level. Data are limited to property value \$100,000-\$1,000,000, Year built between 2000-2019, Lot sizes of 2,000-20,000 sq. ft., and Building area of 800-3,600 sq. ft. They are also limited to GSE borrowers with CLTVs <= 80% since waivers have an eligibility criterion.

- Here we examine the gap within the each group between the Refi human appraisal and the Waiver.
  - Whites with a human appraisal get 0.8% lower value than Whites with a refi waiver and this result is highly statistically significant (at 1% level).
    - Whites serve as the control group because their appraisals should not be biased and the Waiver is in theory color blind.
  - Blacks with a Refi human appraisal get 1.4% lower value than Blacks with a refi waiver appraiser. However, this result is not statistically significant at the 5% level, with a confidence range of +0.3% to -3.2%.
  - The difference between the White and Black groups is 0.6%, basically the same as for the AVM and the Pair approaches.
- This result further supports our conclusion that intentional and unintentional racial bias by appraisers on refinance loans is uncommon and not systemic.

# FHA Purchase Transactions: A Unique Natural Experiment to Evaluate Bias

- FHA purchase appraisals are helpful as a check because they are prone to anchoring in two ways:
  - Sales price as usual.
  - Also prone to LTV anchoring due to the statutory LTV limit of 96.5%.
- There is another helpful characteristic: an average LTV of 95.23%, only slightly below the statutory maximum 96.5%.
- As the chart below demonstrates:
  - FHA purchase loan median LTV equals 96.5% for all listed groups.
    - Since 73% of FHA loans are at the cap, the median LTV equals the cap.
  - FHA purchase loan mean LTVs are all in a tight range relative to median LTVs (White and Black applicants have a 0.48% mean LTV difference).

	LTV of FHA Purchase Loans		
	# of obs.	Median	Mean
White	375,508	96.50	95.04
Black	92,985	96.50	95.52

Note: Data cover an estimated 99% of the agency market.  
Source: AEI Housing Center National Mortgage Risk Index.

# FHA Purchase Transactions: a Natural Experiment to Evaluate Bias (Cont'd)

		Gap btw. Blacks and Whites	# of observations
Purchase	AVM Approach (all loan types)	-0.8%***	299,930
	FHA only	1.0%***	35,070
Refi	AVM Approach (all loan types)	-0.7%***	243,370
	FHA only	-0.4%	30,992

Note: The results are differences in property value relative to Whites. \*\*\* denotes significance at the 1% level. Data for this Approach are limited to property value \$100,000-\$1,000,000, Year built between 2000-2019, Lot sizes of 2,000-20,000 sq. ft., and Building area of 800-3600 sq. ft.

- The results for FHA purchase loan appraisals, which involve substantial numbers of Black applicants, provide additional support for the existence of small appraisal valuation gaps, which cannot be attributed to racial bias (the positive gap for FHA purchase loans to Blacks may be due to the larger seller concessions allowed under the FHA program compared to conventional loans and which tend to be used more in low-income areas. Such concessions are absent on refis).
  - As noted earlier, there is minimal to no difference between White and Black LTV levels.
  - There are two outcomes that would result from bias, both of which are unlikely to occur.
    - A biased lower valuation for Blacks, which if present in practice, would result in the frequent renegotiation of purchase contracts to a lower price, since the value does not support the loan amount and there is little LTV room. However, this would work to the benefit of the home buyer; and the CRN survey indicates renegotiations are infrequent.
    - To display a bias, the appraiser abandons the tendency to anchor to both the purchase price and a key LTV level, yet the dollar difference here is so small that it itself would be red flag and would create a notable headache for the appraiser over the small amount.
- The results for refinance loans are fairly similar to purchase loans, and as noted above, the results for purchase loans more or less rule out significant bias.
- Therefore this experiment also supports our conclusions.

# A Further Robustness Test: Refi Appraisal Gap by Neighborhood

- Under this test, we create census tract groups based on shares of Black-residents and then compare gaps for tracts with  $\leq 25\%$  and  $> 25\%$  Black residents.
- This is an important test as the media allegations of racial bias commonly occurred in predominantly White neighborhoods.
- The gaps for tracts with  $\leq 25\%$  and  $> 25\%$  Black residents are virtually identical ( $-0.6\%$  and  $-0.7\%$  respectively), both of which are nearly the same gap as shown on Slide 9:  $-0.7\%$ ).
- This result adds further support to our conclusion that the alleged practice of knowing the applicant's race results in racial bias by appraisers on refinance loans is uncommon and not systemic.
- It also supports the conclusion that unintentional appraiser bias based on race is also uncommon and not systemic.

Black resident share of census tract	Gap btw. Blacks and Whites	# of White observations	# of Black observations
All	$-0.7\%^{***}$	212,935	30,435
$\leq 25\%$	$-0.6\%^{***}$	205,480	19,639
$> 25\%$	$-0.7\%^{**}$	7,198	10,774

Note: Regression is the AVM refi approach and data are limited to property value \$100,000-\$1,000,000, Year built between 2000-2019, Lot sizes of 2,000-20,000 sq. ft., Building area of 800-3600 sq. ft. We only use human appraisals, waivers are excluded. Minority share of census tract is defined as the share of Black people residing in a tract. **\*\* denotes significance at the 5% level and \*\*\* at the 1% level.**

# Conclusion

We set out to statistically examine the level of racial bias in human performed appraisals. The CRN survey of lenders and AMCs suggest that Reconsiderations of Value (ROVs) are infrequent, as are reappraisals based on an ROV. Further, ROVs with an allegation of racial bias on 1<sup>st</sup> appraisal are also infrequent.

The CRN survey and HMDA data both support the conclusion that our data set, consisting entirely of closed loans, does not suffer from a significant level of selection bias.

Thus, statistically analyzing big data on closed loans can contribute to determining the presence and levels of racial bias by appraisers.

The more biased appraisals there are and the larger the under-valuation, the easier it would be to detect them in the results – and vice versa.

To the extent a claim of disparate impact is made, it does not need to demonstrate intent or knowledge. However, any challenged practice must be shown to be the substantial cause of the disparate impact.

Our statistical analysis looked for evidence of either intentional racial bias, as cited in media reports, or unintentional.

Our conclusions are (i) media allegations of appraiser racial bias on refinance loans is uncommon and not systemic and (ii) a claim of unintentional bias on refinance loans, if the basis of a disparate impact claim, was also found to be uncommon and not systemic.