

MAY 2022

Racial & Ethnic Valuation Gaps in Home Purchase Appraisals - A Modeling Approach

Executive summary

This Research Note builds on our previous Note published in September 2021¹ to report on a refined and expanded modeling approach that Freddie Mac adopted in the racial valuation gap context. Our first Note analyzed tract averages as little research existed on this important topic and we wanted to spark a discussion and receive feedback on our approach leading to subsequent work. This note reflects much of the feedback we have received.² Specifically we adopt a modeling approach to study whether properties in predominantly minority (Black and Latino) census tracts are more likely than properties in predominantly White census tracts³ to receive an appraisal value that is lower than the contract price⁴ when homes are appraised for the purpose of obtaining a home loan (home purchase appraisals.⁵) Our model results indicate that, even after controlling for important factors that affect house values and appraisal practices, appraisal outcomes still differ for properties in predominantly Black and Latino tracts relative to those in predominantly White tracts.

This new effort controls for a wider variety of variables, including house characteristics, neighborhood characteristics, housing market dynamics, and fixed effects. It applies multiple modeling techniques to examine three questions: (1) whether properties in minority tracts are more likely to receive “appraisal

-
- 1 Link to the first Research Note: [Racial and Ethnic Valuation Gaps in Home Purchase Appraisals](#).
 - 2 We want to thank all the researchers who have provided thoughtful and constructive feedback. Although researchers may have different perspectives in terms of the detailed modeling approach, we appreciate the suggestions and dialogue that help advance the continued research on this important topic.
 - 3 We conduct analyses at the tract level for two reasons: 1) we are assessing whether the racial or ethnic composition of the neighborhood affects appraisal valuation, and 2) there is no data on buyers or sellers’ race or ethnicity in appraisals. When individual race and ethnicity data are not available, it is a common practice to use minority share at the tract level to classify tracts and analyze minority tracts. Considering the racial or ethnic composition of the neighborhood in appraisal valuation is an illegal form of discrimination regardless of the race or ethnicity of the buyer or seller. This method meets the objective of this research since it enables researchers to examine the neighborhood-level differences between predominantly minority tracts and predominantly White tracts. It is also interesting to study whether and how buyer (or seller) race/ethnicity plays a role on appraisals, conditional on the racial makeup of the tract. We consider this as an important potential future research task.
 - 4 For the purposes of this research, an “appraisal gap” means the percent difference between minority and White tracts in the share of properties receiving appraised values that are lower than the contract price for home purchases. This measure is called “appraisal value lower than contract price.” We acknowledge that the sale price is not always equal to market value, and we expect that in all areas some appraisals will report values lower than the contract price. However, research data indicate that a high percentage of appraisals are at or above the purchase contract price (Calem, Lambie-Hanson, and Nakamura, 2017).
 - 5 While others have compared appraisal estimates to automated valuation (AVM) estimates, we believe this could confound certain analysis. AVM estimates are statistically based computer programs that use real estate information such as comparable sales, property characteristics, and price trends to provide a current estimate of market value for a specific property. Appraisal estimates may come below the AVM estimate if the AVM estimate is too high because of issues with modeling instead of unobserved effects like potential socioeconomic or racial bias. Those conversations then steer in a different direction debating the statistical properties of modeled AVM estimates rather than prices agreed upon by willing buyers and willing sellers. With those cautions in mind, we leverage our assumption that market-oriented valuations and negotiations lead to a more appropriate signal and thus compare appraisal estimates of value to contract price. Using price as our ‘north star’ limits our comparison, though, to purchase transactions as contractually agreed upon prices by two parties do not exist for refinance transactions. We acknowledge and recognize that recent studies have pursued a different focus because they tried to eliminate potential bias of at least one of those participating parties, but such was not our approach.



value lower than contract price” compared to those in White tracts;⁶ (2) whether the likelihood for homes in minority tracts to receive “appraisal value lower than contract price” increases as the minority concentration increases in the tract; and (3) whether houses in minority tracts receive even lower appraisal values than those in White tracts in instances of “appraisal value lower than contract price”.

The purpose of this Research Note is to share the modeling details, including data sources, variables selected for the model estimation, the conceptual framework, model specifications (described in sections 1 and 2) and our findings (discussed in section 3). After controlling for important factors that affect house values and appraisal practices, our research shows the following outcomes: (1) properties in Black and Latino tracts are more likely to receive “appraisal value lower than contract price” than those in White tracts; (2) the likelihood for properties in Black tracts and Latino tracts to receive “appraisal value lower than contract price” increases as the Black or Latino concentration increases in the tracts; and (3) when “appraisal value lower than contract price” takes place, houses in Black tracts are appraised slightly lower relative to comparable homes in White tracts in terms of the percent difference between the appraisal value and the contract price. We conclude the paper by discussing the implications for future research and policy. Our modeling findings shed light on the drivers for the appraisal differences observed in our dataset; however, additional research is required to identify the causal mechanisms behind differences in appraisal outcomes. Moreover, to develop appropriate solutions, there is a need to further investigate the impact of varying appraisal outcomes on buyers and sellers, as well as the communities they live in, especially in cases of rapidly changing home prices.

Introduction

It is possible that the pre-modeling appraisal gaps we observed in our first Research Note⁷ are attributable to variations in the characteristics of the properties or neighborhoods. Accordingly, we examine this possibility through a rigorous modeling approach by isolating the effect of racial and ethnic appraisal differences for comparable homes. Through the modeling journey, we aim to answer three research questions:

1. Does the minority tract flag⁸ help explain appraisal gaps after controlling for important factors that affect house values and appraisal practices?
2. Does the minority concentration in tracts help explain appraisal gaps after controlling for important factors that affect house values and appraisal practices?
3. Does the minority tract flag help explain how much lower the appraisal value is relative to the contract price when “appraisal value lower than contract price” happens after controlling for important factors that affect house values and appraisal practices?

6 Throughout this Research Note, the term “minority” refers to Black or Latino, while the term “White” refers to White non-Latino. This study compares tracts where Blacks comprise the majority of residents (so-called “Black tracts”) or Latinos comprise the majority of residents (so-called “Latino tracts”) to tracts where White non-Latino residents are the majority (“White tracts”). “Tracts” refer to census tracts. They are small subdivisions within a county typically containing between 1,200 and 8,000 people. Census tract boundaries are likely not consistent with actual neighborhood boundaries as recognized by market participants; however, for purposes of this research, we use the terms “neighborhood” and “census tracts” interchangeably.

7 We acknowledge AEI for their related research efforts on this topic. Notably, the pre-modeling appraisal gaps in Freddie Mac’s first Research Note align to those in the AEI analysis (Pinto & Peter, 2021). For example, the Freddie Mac Note reported an appraisal gap of 5.3% in Black tracts compared to 5.2% in the AEI analysis.

8 The flag indicates whether the majority residents in the tract are Black (Latino).



1. Data

Our primary data source includes over 12 million Single Family one-unit appraisals for purchase transactions submitted to Freddie Mac from January 1, 2015 to December 31, 2020 through the Uniform Collateral Data Portal (UCDP). This dataset covers all CBSAs⁹ and more than 95% of census tracts across the country. When there are multiple submissions in the UCDP, the first appraisal is used. We exclude non-arm-length transactions (i.e., a non-arms-length sale is a sale between related parties), contracts with a concession amount exceeding 3%, and appraisal values or property features outside the normal range (for example, properties with more than ten bedrooms or more than four stories). In addition, some data outliers are excluded, such as properties with unacceptable conditions per Freddie Mac’s underwriting standards.

UCDP provides data on property characteristics we need for this research. The neighborhood characteristics are collected from American Community Survey (ACS) and Home Mortgage Disclosure Act (HMDA) data.¹⁰ The ACS 2015-2019 five-year estimates at the tract level are used to match our research period. The HMDA data provides valuable data points for national housing market dynamics.¹¹ The rest of the section explains how the dependent variables and explanatory variables are selected for the three research questions.

Dependent variables

We constructed the dependent variables to match the research objectives. For the first and second research questions, the dependent variable captures the incidence of “appraisal value lower than contract price” using an indicator variable equal to

- > 1 if the appraisal value comes below the contract price, and
- > 0 if the appraisal value matches or exceeds the contract price.

For the third research question, the dependent variable captures the degree to which the appraised value is less than the contract price. It is expressed as the percent difference between the appraisal value and the contract price, calculated as $\frac{\text{Appraisal Value} - \text{Contract Price}}{\text{Contract Price}} \times 100\%$.

9 CBSA stands for Core-Based Statistical Areas. Metropolitan and Micropolitan Statistical Areas are collectively referred to as Core-Based Statistical Areas (CBSAs). More details on CBSA can be found at: <https://www.census.gov/topics/housing/housing-patterns/about/core-based-statistical-areas.html>. In modeling, “control for appraiser-CBSA fixed effects” means creating dummy variables for individual appraisers at each CBSA and including them in the model estimation process.

10 We have also explored neighborhood characteristics at the county level based on references provided by Perry et al. (2018). The County Business Pattern (CBP) data is an annual series that provides subnational economic data by industry. We could use the CBP 2019 to get data on essential establishments in counties, such as food or drink places and gas stations. However, these county-level variables are not ideal for tract-level studies since big differences between certain census tracts could exist within the same county. Considering this limitation and the other variables already included in our models, we decided not to include county-level variables from CBP in our final models. As a robustness check, we have estimated our models with these county-level factors and find our modeling conclusions continue to hold.

11 The Home Mortgage Disclosure Act (HMDA) requires many financial institutions to maintain, report, and publicly disclose loan-level information about mortgages. The public data are modified to protect applicant and borrower privacy. The key fields in the public data include dozens of fields such as loan purpose and loan amount. We use HMDA data to construct the turn-over rate of purchase transactions at the tract level, as explained in the “Data” section. More details on HMDA data can be found at: <https://www.consumerfinance.gov/data-research/hmda/>.



Explanatory variables

For the first and the third research questions, our main explanatory variable of interest is the so-called minority tract flag. We classify tracts according to the minority share of the population in that tract based on the 2010 census data and use 50% as the threshold. Specifically, if the share of Black (Latino) residents in a tract is 50% or higher, this tract is categorized as Black (Latino). If the overall minority share¹² in a tract is below 50%, the tract is flagged as White. When comparing Black (Latino) tracts to White tracts, the flag is 1 for properties that are in Black (Latino) tracts and 0 for those in White tracts. For the second research question, we focus on the minority concentration in tracts as a continuous explanatory variable, which is the share of the Black (Latino) population in the tract where the property is located.¹³

The remainder of the explanatory variables are carefully selected to control for confounding factors that may influence home values and shape appraisal practices. They are employed to address all three research questions and can be grouped into four major categories: house characteristics, neighborhood characteristics, housing market dynamics, and fixed effects.

Under the category of house characteristics, we select property features based on the widely cited hedonic model literature, which uses a variety of house characteristics to predict house value.¹⁴ Our rich set of house characteristics sourced from UCDP includes living area square feet, indicators of fireplace, pool and garage, number of baths, number of bedrooms, number of stories, and construction year.¹⁵

Our neighborhood characteristics come from multiple sources. At the census tract level, we control for neighborhood demographics with the following variables: median household income, median age, share of households with child(ren), share of population that is in the labor force, and share of occupied housing units that are owner-occupied. These features are collected from ACS 2015-2019. We also consider the urbanization level by building an urban, suburban vs. rural flag.¹⁶

In addition, we develop three important measures to control for housing market dynamics that matter for appraisal practices. The first one is a turn-over rate of purchase transactions at the tract level. It is calculated as the average annual purchase transactions per square mile based on HMDA data. This variable approximates the availability of potential comparable sales (comps¹⁷) for the subject property being appraised. The second measure is a gentrification flag. This factor is important because houses in gentrified areas are usually harder to appraise due to rapid growing housing prices and increasing heterogeneity of houses in terms of size, quality, and condition. We define a tract as undergoing gentrification if it was low-income in the beginning data period and experienced a rapid income growth by the end of the period.¹⁸ The third one is the Forecast Standard Deviation (FSD), which is generated

12 Anyone who is not White non-Latino is counted in the overall minority group.

13 When minority tract flags are used as explanatory variables, we estimate the models for Black and Latino separately. When examining the difference between Black and White tracts, we use only observations in Black tracts and White tracts. Similarly, when we study the difference between Latino and White tracts, we include only observations in Latino tracts and White tracts. This approach ensures that the model's baseline is only White tracts. Another reason for the separate model estimation is that Black tracts and Latino tracts are not mutually exclusive. Some tracts are both Black and Latino even though the fraction of these tracts is small. On the other hand, when Black and Latino shares are used as explanatory variables, we include all observations for model estimation.

14 For a theoretical presentation of hedonic models, refer to Rosen (1974) and Muellbauer (1974).

15 We decided not to include controls that may involve appraisers' discretion, including property conditions, construction quality and views. However, our modeling results barely change when those variables are included.

16 We classify urban/suburban/rural using the census-convenient method as described in Airgood-Obyrcki and Rieger (2019).

17 In a residential appraisal, the value is developed primarily by examining other competitive homes that have been sold recently. Those sales are called "comparable sales," commonly called "comps."

18 A tract is low-income if its median income ratio to the state average is less than 70%, and it is growing quickly if the income ratio increased by more than 10 percentage points. Our definition is borrowed from Neal et al. (2020) and Gould Ellen and O'Regan (2008). There is no commonly accepted definition of gentrification. Alternatives can be found in papers such as Brummet and Reed (2019) and Richardson et al. (2019).



by Freddie Mac’s internal automated valuation model (AVM). It captures the level of uncertainty when predicting a house’s value. The larger the FSD, the more uncertain the model is about the estimated value.¹⁹

Furthermore, we control for appraiser and spatial heterogeneity through appraiser-CBSA fixed effects. The inclusion of appraiser-CBSA fixed effects accounts for appraiser-level heterogeneity. For example, some appraisers may prefer algorithms while others use more traditional methods, or some appraisers do business primarily in gentrified areas while others in stable markets. It also handles spatial heterogeneity between Black (Latino) tracts and White tracts that may not be fully captured by neighborhood characteristics. Year and month dummies are also included in the model to capture changes in the housing markets over time, such as house price appreciation and COVID-19 shocks. Lastly, we include an FHA appraisal indicator from the UCDP dataset to control for potential overbidding in the contract price negotiation process.²⁰

2. Conceptual framework

To answer the three research questions, we leverage both traditional econometric and machine learning (ML) techniques. In the traditional econometric model, we built Ordinary Least Square (OLS) and Logistic Regression (logit) models. In the ML path, we tested various models and Light Gradient Boosted Machine (LightGBM) performed the best.²¹ This section presents our detailed modeling framework and the associated rationale.

The OLS model specification is as follows:

$$Y_{int} = \alpha \text{Minority}_n + X_i\beta + W_n\gamma + \mu_{jc} + \theta_t + \mu_m + \epsilon_{int}$$

The dependent variable Y_{int} is a dummy indicating “appraisal value lower than contract price” (for the first and second research questions) or the percent difference between the appraisal value and the contract price (for the third research question) for house i in census tract n in year t . The explanatory variable Minority_n refers to the minority tract flag (for the first and third research questions) or the minority share in the tract where the associated property is located (for the second research question). As detailed in the “Data” section, on the right-hand side we control for house characteristics (X_i),

19 Although it would be ideal to control for house price index (HPI) at the tract, we decided not to do so because of data availability. HPI is constructed based on repeated sales. If a census tract does not have enough repeated sales in the data period, the HPI will not be reliable. We have done various robustness checks to ensure that this decision does not impact our modeling conclusions. First, we test by using the census-tract level HPI downloaded from FHFA. Second, we use the county-level HPI based on Freddie Mac’s weighted-repeat-sales index (WRSI). Third, we estimate the model with appraiser-CBSA-year fixed effects to fully control for all time-variant confounding factors at the CBSA level. Based on the above robustness checks, the OLS coefficients barely change.

20 We have done the robustness check by excluding FHA appraisals and found that our modeling conclusions stay unchanged. Our decision is to keep the FHA appraisals in our modeling dataset by adding an FHA indicator. This enables us to employ a similar dataset as used in the first research note. Please note that this variable is different from the share of FHA loans at the tract level. The recently published AEI research (2021) states that FHA loan borrowers are more likely to be inexperienced first-time home buyers and are more likely to overbid. Their research uses the share of FHA loans at the tract level based on the 2020 HMDA data. Our internal research suggests that this variable is much more predictive of Black tracts than it is of “appraisal value lower than contract price”, and thus it can be a potential proxy variable for Black neighborhoods. Therefore, we decided to not include this variable. We also performed a robustness check by including this potential proxy in our models and found that our major modeling conclusions hold well. Instead of using the share of FHA loans at the tract level, we include an FHA appraisal indicator in our models. Similarly, we intentionally chose not to include borrower characteristics such as average credit score or share of one-borrower loans at the tract level. These factors can be proxies for race or ethnicity, which we are trying to measure in our regression results. Also, appraisers are not supposed to consider these factors during the appraisal process and thus we want our statistical estimations to align with the practitioners’ approach.

21 The Light Gradient Boosting Machine (LightGBM) is a framework for machine learning that is based on decision tree algorithms and used for ranking, classification, and other machine learning tasks.



neighborhood characteristics and market dynamics (W_n), appraiser-CBSA fixed effects (μ_{jc})²², and year and month dummies (θ_t and μ_m). The error term is denoted by ϵ_{int} .

The equation that follows shows the logit model specification. Instead of assuming a linear specification between the dependent and explanatory variables, the logit model specifies a linear specification between the log odds ratio and explanatory variables. In the model, p_{int} refers to the probability of receiving “appraisal value lower than contract price.”²³ Variables on the right-hand side have similar meanings as their counterparts in the OLS model.

$$\log\left(\frac{p_{int}}{1-p_{int}}\right) = \alpha \text{Minority}_n + X_i\beta + W_n\gamma + \mu_j + \tau_c + \theta_t + \mu_m + \epsilon_{int}$$

Although it is common practice to estimate a logit model for dependent variables with binary outcomes, we build OLS models for three main reasons. First, it is easier to interpret OLS results. In our OLS model, the estimator α is equal to the difference in the probability of Black (Latino) and White tracts receiving “appraisal value lower than contract price.” In contrast, the estimator α from the logit model reflects only the difference in the log odds ratio. Second, a linear probability model is free from incidental parameter bias that is generated by the inclusion of fixed effects.²⁴ Third, OLS is efficient in runtime as it differences out fixed effects,²⁵ which is why OLS does not suffer from incidental parameter bias.

To further strengthen our modeling results we employ advanced ML techniques in addition to traditional econometrics. Compared with linear models, ML models such as XGB and Light GBM models typically have stronger predictive power because they are more robust to data outliers, take advantage of non-linear relationships when needed, and thus can fit estimation data more accurately. A variety of model agnostic methods have been developed to interpret ML results, especially the impact of model features on the outcome. In this research, we build LightGBM models and use partial dependence plots (PDP) to interpret the impact of minority tract flags or minority concentration on the outcome.

To further strengthen our modeling results we employ advanced ML techniques in addition to traditional econometrics. Compared with linear models, ML models typically have stronger predictive power.

22 We use the interacted appraiser-CBSA fixed effect for OLS instead of separate two-way fixed effects for two reasons: 1) the combination allows us to control for potential variation in appraisers’ practice in different CBSAs, and 2) as shown in our data, appraisers typically work in only one or two CBSAs, which makes appraiser dummies strong indicators for location. In addition, we estimate the model using various ways to construct the fixed effects. First, we run two-way fixed effects (appraiser and CBSA). Second, we test appraiser-CBSA-year combinations. Third, in the machine learning (ML) models, we allow fully non-parametric combinations of fixed effects. Based on these robustness checks, our main modeling results hold well.

23 For more details on the interpretation of a linear probability model estimator, see Angrist and Pischke (2008).

24 For a discussion of incidental parameter bias, see Lancaster (2000).

25 Our data include more than 52,000 appraisers and more than 120,000 appraiser-CBSA combinations.



3. Modeling results

This section presents our findings to the three research questions posed earlier.

Does the minority tract flag²⁶ help explain appraisal gaps after controlling for important factors?

After controlling for important factors that affect house values and appraisal practices, we find that properties in Black and Latino tracts are more likely to receive “appraisal value lower than contract price” than those in White tracts.

The OLS results²⁷ presented in **Exhibit 1**²⁸ show that properties in Black and Latino tracts are more likely to receive “appraisal value lower than contract price” relative to those in White tracts after controlling for important factors (for the full modeling results for Exhibit 1, see Appendix 1). The first three columns in Exhibit 1 show the step-by-step regression results for Black tracts and the last three columns show the step-by-step results for Latino tracts. The step-by-step results demonstrate how coefficients of the Black (Latino) tract flag change as more controls are added to the model. In column one and column four, when no controls are included in the model, the magnitude of the coefficients on minority tract in these regressions is close to the raw difference we observed in the first Research Note.²⁹ In column two and column five, the results are presented for regressions with controls (house characteristics, neighborhood characteristics, housing market dynamics, and FHA appraisal indicator)³⁰ and time trends (year and month dummies). The coefficients attenuate compared to the raw comparison in column one and four.

EXHIBIT 1:

Ordinary Least Squares (OLS) results for the first research question

	Depvar: Appraisal Lower Than Price					
	Black			Latino		
Black tract flag	0.049*** (0.009)	0.042*** (0.008)	0.024*** (0.003)			
Latino tract flag				0.080*** (0.014)	0.062*** (0.014)	0.029*** (0.003)
House & neighborhood controls	No	Yes	Yes	No	Yes	Yes
Year & month dummies	No	Yes	Yes	No	Yes	Yes
Appraiser-CBSA fixed effects	No	No	Yes	No	No	Yes
Depvar mean of white tracts	0.073	0.073	0.073	0.073	0.073	0.073
Adjusted R ²	0.001	0.010	0.076	0.004	0.014	0.079
N	10899546	10898846	10698843	11081215	11080594	10878762

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

26 The minority tract flag is defined based on the minority concentration at the tract level; therefore, the modeling results should be interpreted at the tract level. Please note that these results are not at the borrower level.
 27 Throughout the OLS modeling process, standard errors are heteroskedasticity-consistent and clustered at the county level.
 28 The number of observations decreases as more controls are added into the regression and observations with missing values are excluded. The second OLS regression is based on a smaller sample size than the first, parsimonious one due to missing value in controls (mainly from log median household income). The third OLS regression is based on an even smaller sample size because appraisers without names are excluded.
 29 The small change in magnitude is due to the exclusion of observations with outlier or missing values.
 30 For the rest of OLS tables, “controls” refer to house characteristics, neighborhood characteristics, housing market dynamics, and the FHA appraisal indicator. This is also applicable for logit tables.



In column three and column six when all the important factors (as explained in the “Explanatory Variables” section) are taken into consideration, the coefficients decline to 2.4% for the Black tract flag and 2.9% for the Latino tract flag.³¹ This translates to a 2.4% higher likelihood of “appraisal value lower than contract price” for the Black tracts compared to White tracts and 2.9% higher likelihood for Latino tracts. Both coefficients are statistically significant at the 0.01 level, meaning that the minority tract flag helps explain appraisal gaps even after controlling for important factors. The magnitudes of the coefficients are both statistically and economically meaningful given the fact that about 7.3% of the houses in White tracts receive “appraisal value lower than contract price.”³²

To check whether the patterns in Exhibit 1 are consistent and robust, we perform several robustness checks from different perspectives. First, we re-estimate the model by constraining the sample to tracts where the concentration of White people is 50% or above. The main purpose of only using White-prevalent tracts is to further control for unobserved heterogeneity between Black (Latino) tracts and White tracts in our original analysis. In the new regressions, we use 30% as the threshold to define minority tracts. The coefficients for the Black and Latino tract flags attenuate a little compared to those reported in Exhibit 1, but they are still statistically significant. Second, we re-estimate the model by using only appraisals submitted by appraisers who did enough work in both Black (Latino) tracts and White tracts.³³ The coefficients for Black and Latino tract flags turn out to be bigger and statistically significant. Third, we divide our estimation dataset into six categories based on contract price and re-estimate the model for each category. The coefficients for the Black and Latino tract flags stay positive and significant in all six price categories. Fourth, we divide our observations into five groups by median household income at the tract level and re-run the model for each subgroup. Our results are persistent across almost all income categories.³⁴ The third and fourth robustness checks confirm that house price and household purchasing power are not likely to be the main driver for appraisal gaps. Fifth, we estimate the model for top 10 CBSAs respectively and observe consistently significant appraisal gaps in those metropolitan cities when there are enough appraisals from minority tracts. Last, we control for important location-specific market dynamics in different ways, including census tract-level house price index³⁵ on the right-hand and using appraiser-CBSA-year fixed effects. These alternative model specifications barely change the coefficients of the minority tract flags and they remain statistically significant.

The minority tract flag helps explain appraisal gaps even after controlling for important factors. Findings indicated a 2.4% higher likelihood of “appraisal value lower than contract price” for the Black tracts compared to White tracts, and 2.9% higher likelihood for Latino tracts.

31 Both “2.4%” and “2.9%” refer to a percentage point increase. Throughout this Research Note, when interpreting OLS and ML modeling results, we use “%” as a convenient, readily understood term that is interchangeable with “percentage point.” One exception happens when interpreting log odds in the context of logit model results. In this instance, “% increase” refers to “percentage increase.”

32 This means that the likelihood increases by 33% ($\frac{2.4\%}{7.3\%}$) for Black tracts and by 40% ($\frac{2.9\%}{7.3\%}$) for Latino tracts.

33 There are 934 appraisers who have a sufficiently large sample for the Black versus White t-test; there are 1,560 appraisers with an adequate sample for the Latino versus White t-test. Please refer to our first Research Note for a detailed description of how these appraisers are selected.

34 We group the sample based on quintiles of median household income at the tract level. The coefficients for the Black tract flag vary from 1.6% to 3.4%, and they are all significant at the 0.01 level. The coefficients for the Latino tract flag vary from 2.3% to 3.1%, and they are all significant at the 0.01 level with one exception in the 5th quintile. This exception is likely caused by the fact that there are few Latino tracts in the top quintile category.

35 The census tract-level house price index (HPI) is downloaded from the FHFA website. This dataset does not cover all the tracts and therefore the robustness check is based on the observations that have HPI from the FHFA datasets.



Exhibit 2 reports the estimated log odds for Black and Latino tract flags (for the full modeling results for Exhibit 2, see Appendix 2).³⁶ Row two shows that the coefficients for Black and Latino tract flags are significant with positive signs after controlling for property characteristics, neighborhood characteristics, housing market dynamics, year and month dummies, and CBSA fixed effects. Row three presents the estimation results based on the subgroup of appraisers.³⁷ The odds ratio for the Black tract flag is 1.519, which implies that the odds of receiving “appraisal value lower than contract price” for properties in a Black tract is 51.9% higher than those in a White tract. Similarly, the odds of receiving “appraisal value lower than contract price” for properties in a Latino tract is 30% higher than those in a White tract.

EXHIBIT 2:

Logit results for the first research question

Row	Estimation	Dataset	Black vs. White				Latino vs. White			
			Coefficient Estimates	Odds Ratio Estimates	P-Value	C-Statistic	Coefficient Estimates	Odds Ratio Estimates	P-Value	C-Statistic
1	No control	Full	0.568	1.764	<0.0001	0.511	0.827	2.286	<0.0001	0.526
2	Controls, year and month dummies, and CBSA fixed effects	Full	0.299	1.348	<0.0001	0.658	0.319	1.376	<0.0001	0.664
3	Controls, year and month dummies, appraiser and CBSA fixed effects	Subset*	0.418	1.519	<0.0001	0.724	0.262	1.300	<0.0001	0.712

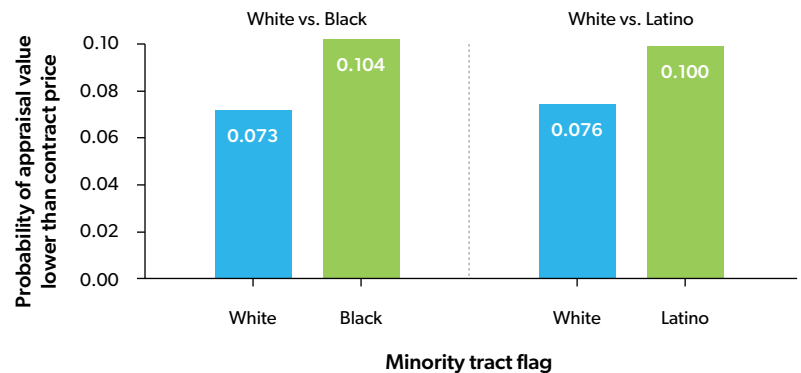
* Note: based on properties appraised by the subgroup of appraisers who have done enough work in both Black/Latino tracts and White tracts.

Lastly, we relax the linearity assumption on model specification via ML techniques. The LightGBM model allows a fully non-parametric relationship between dependent and explanatory variables. The Area Under the Curve (AUC) for the Black (Latino) versus White regression is 0.719 (0.722). The partial dependence plots (PDPs) in **Exhibit 3** show that properties in Black (Latino) tracts are 3.1% (2.4%) more likely to get “appraisal value lower than contract price.” These results are consistent with those from the OLS and logit models.

EXHIBIT 3:

Machine Learning (ML) results for the first research question

Partial dependence plot for probability of appraisal value lower than contract price by Black or Latino tract flag.



36 Due to computational power limitations, we include appraiser and CBSA fixed effects only when estimating the model based on a subset of the full data.

37 The same subset is used for the robustness check of the OLS results.



Does the minority concentration in tracts help explain appraisal gaps after controlling for important factors?

After controlling for important factors that affect house values and appraisal practices, we find that the likelihood for homes in Black (Latino) tracts to receive “appraisal value lower than contract price” rises as the Black (Latino) concentration increases in the neighborhood.

To further explore the relationship between minority neighborhoods and the likelihood of receiving “appraisal value lower than contract price,” we estimate alternative models using the Black share and Latino share (instead of the Black or Latino tract flag) as the explanatory variables. The alternative OLS results presented in **Exhibit 4** show that a 1% increase in Black population corresponds to an increase of likelihood by 0.08% before controls and an increase of likelihood by 0.05% after full controls. Similarly, a 1% increase in Latino population corresponds to an increase of likelihood by 0.15% before controls and an increase of likelihood by 0.06% after full controls. All the coefficients are statistically significant at the 0.01 level (for the full modeling results for Exhibit 4, see Appendix 3).

Consider an example to interpret the results. Suppose the Black concentration in a Black tract is 75% and the Black concentration in a White tract is 25%. Then the likelihood of receiving an “appraisal value lower than contract price” increases by $0.05\% \times \left(\frac{75\% - 25\%}{1\%}\right) = 2.5\%$ for the Black tract. This is close to the 2.4% coefficient for the Black tract flag as reported in Exhibit 1.

EXHIBIT 4:

Ordinary Least Squares (OLS) results for the second research question

	Depvar: appraisal lower than price		
Black share (%)	0.0008*** (0.0001)	0.0009*** (0.0001)	0.0005*** (0.0001)
Latino share (%)	0.0015*** (0.0002)	0.0014*** (0.0002)	0.0006*** (0.0000)
House & neighborhood controls	No	Yes	Yes
Year & month dummies	No	Yes	Yes
Appraiser-CBSA fixed effects	No	No	Yes
Depvar mean	0.083	0.083	0.083
Adjusted R ²	0.010	0.019	0.083
N	12623454	12622663	12406292

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

EXHIBIT 5:

Logit results for the second research question

Row	Estimation	Dataset	Black share			Latino share			C-statistic
			Coefficient estimates	Odds ratio estimates	P-value	Coefficient estimates	Odds ratio estimates	P-value	
1	No control	Full	0.010	1.010	<0.0001	0.016	1.016	<0.0001	0.595
2	Controls, year and month dummies, and CBSA fixed effects	Full	0.006	1.006	<0.0001	0.008	1.008	<0.0001	0.670
3	Controls, year and month dummies, appraiser and CBSA fixed effects	Subset*	0.008	1.008	<0.0001	0.007	1.007	<0.0001	0.716

* Note: based on properties appraised by the subgroup of appraisers who have done enough work in both Black/Latino tracts and White tracts.

Exhibit 5 presents the alternative logit model results, which show that a higher Black or Latino concentration is associated with a higher chance of receiving “appraisal value lower than contract price” (for the full modeling results for Exhibit 5, see Appendix 4). This pattern holds before and after controls, and all coefficients are statistically significant. The results in row three are based on the appraisals



submitted by appraisers who did enough work in both Black (Latino) tracts and White tracts.³⁸ The odds ratio for the Black share is 1.008, which implies that for every 1% increase in Black concentration, there will be an increase of 0.8% in the odds of receiving “appraisal value lower than contract price.” Similarly, the odds ratio for the Latino share is 1.007, which implies that for every 1% increase in Latino concentration, there will be an increase of 0.7% in the odds of receiving “appraisal value lower than contract price.”

Our machine learning model yields similar findings. The partial dependence plots (PDPs) in **Exhibit 6** show a clear trend that the probability of receiving an “appraisal value lower than contract price” increases as the percentage of Black or Latino people increases. For Black tracts, the probability goes up by about 3% when the percentage of Black population in a tract increases from 0% to 50%. For Latino tracts, the likelihood increases by approximately 3.5% when Latino concentration increases from 0% to 60%.

To summarize the modeling results so far, both the minority tract flag (0/1) and the minority concentration (%) can explain appraisal gaps after controlling for important factors. As defined earlier, appraisal gaps in our research refer to the different likelihood of receiving “appraisal value lower than contract price” between minority tracts and White tracts. After observing these modeling results, we decided to expand our research by including the “severity” component in addition to “likelihood.” This analysis is reported in the following section.

After controlling for important factors, does the minority tract flag help explain how much lower the appraisal value is relative to the contract price when “appraisal value lower than contract price” happens?

When “appraisal value lower than contract price” takes place, we find that houses in Black tracts are appraised slightly lower than those in White tracts after controlling for important factors that affect house values and appraisal practices.

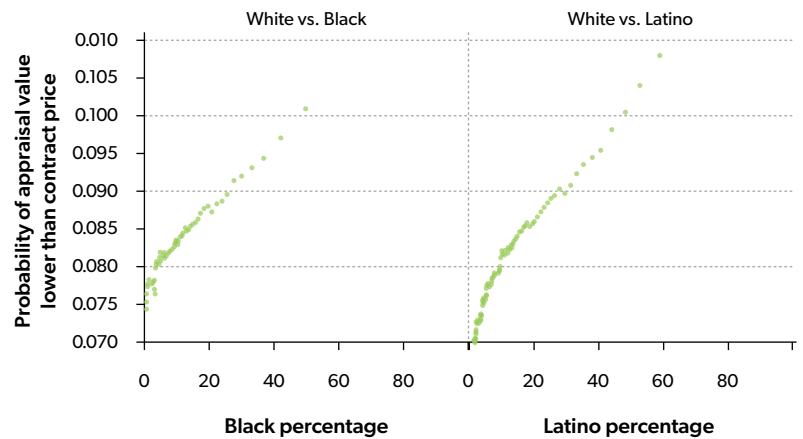
Both likelihood and severity are important in measuring the magnitude of the different appraisal outcomes for minority groups. To explore whether properties in minority tracts experience different severities when “appraisal value lower than contract price” happens,³⁹ we use the same set of explanatory variables for answering the first and second research questions, but use the percent difference between the appraisal value and the contract price as the dependent variable, i.e.,

$$\frac{\text{Appraisal Value} - \text{Contract Price}}{\text{Contract Price}} \times 100\%.$$

EXHIBIT 6:

Machine Learning (ML) results for the second research question

Partial dependence plot for probability of appraisal value lower than contract price by Black or Latino concentration.



38 The same subset is used in Exhibit 2.

39 The estimation sample for the third research question is restricted to properties that receive appraisals below the contract price., i.e., the dependent variable is below zero.



Exhibit 7 displays the OLS results before and after controlling for important factors (for the full modeling results for Exhibit 7, see Appendix 5). The first (fourth) column shows the raw severity differences between Black (Latino) tracts and White tracts.⁴⁰ When appraisal values fall below contract prices, properties in White tracts are on average appraised 4.4% lower than the contract price, while those in Black (Latino) tracts are valued even lower by an additional 1.4% (0.3%). That is to say, the severities for properties in Black and Latino tracts are approximately 5.8% and 4.7%, respectively. These severity levels are comparable to those reported by recent literature (Pinto and Peter, 2021).

When controlling for important factors that can affect home values and appraisal practices, the severity difference between Black and White tract decreases to 0.74%, but it is still statistically significant at the 0.01 level. However, different modeling results are observed for Latino tracts. As control factors are added to the model gradually, the coefficient for the Latino tract flag changes from -0.319% (significant at 0.01 level) to -0.179% (significant at 0.1 level) and then to -0.183% (significant at 0.01 level), which signals unstable modeling results.

Our ML models yield similar results to the OLS modeling. As shown by the partial dependence plots (PDPs) in **Exhibit 8**, properties in Black tracts are appraised slightly lower by about 0.7% (0.052–0.045=0.007=0.7%) relative to those in White tracts when appraisal values fall below the contract price. When comparing Latino tracts to White tracts, the difference is not noticeable (0.043–0.043 = 0).

According to our modeling results, properties in Black tracts are valued slightly lower by about 0.7% (based on both OLS and ML models) relative to those in White tracts when their appraisal values fall below the contract price; however, there is no significant difference between Latino

EXHIBIT 7:

OLS results for the third research question

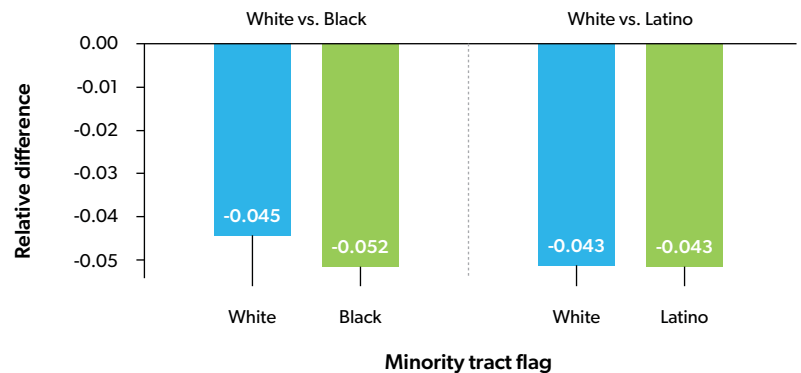
	Depvar: (appraisal – price)/price x 100%					
	Black			Latino		
Black tract flag	-1.435*** (0.137)	-0.900*** (0.082)	-0.741*** (0.052)			
Latino tract flag				-0.319*** (0.103)	-0.179* (0.096)	-0.183*** (0.062)
House & neighborhood controls	No	Yes	Yes	No	Yes	Yes
Year & month dummies	No	Yes	Yes	No	Yes	Yes
Appraiser-CBSA fixed effects	No	No	Yes	No	No	Yes
Depvar mean of white tracts	-4.431	-4.431	-4.431	-4.431	-4.431	-4.431
Adjusted R ²	0.008	0.071	0.186	0.001	0.064	0.179
N	816338	816289	790162	855596	855556	828911

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

EXHIBIT 8:

Machine Learning (ML) results for the third research question

Partial dependence plot for relative difference between appraised value and sales price by Black or Latino tract flag.



40 In addition to the raw severity differences reported in this section, we also examine the raw severity difference by house value categories for properties in Black vs. White tracts. Our analysis shows that when appraised values are lower than the contract price, (1) properties in Black tracts are valued lower than those in White tracts when their values are below \$500,000, and (2) the percent differences are larger for properties in the lower value buckets.



and White tracts. To put the severity difference for Black tracts into perspective, we dollarize the impact from this severity gap. Based on our dataset, when “appraisal value lower than contract price” takes place, the average contract price of these properties in Black tracts is \$243,343; thus, the estimated dollarized impact is \$1,703 according to both the OLS model and the ML model ($\$243,343 \times 0.7\%$).

4. Discussion and implications for further research

In this research note, we examine racial and ethnic valuation gaps in home purchase appraisals through a modeling approach. The traditional econometric and machine learning models we deploy offer very similar conclusions.⁴¹ Based on our modeling results, even after controlling for important factors that affect house values and appraisal practices,⁴² properties in Black and Latino tracts are more likely to receive appraisal values that fall below contract prices, and this likelihood increases as the Black or Latino concentration in the neighborhood increases. In addition, when “appraisal value lower than contract price” takes place, houses in Black tracts are appraised slightly lower relative to those in White tracts, while the severity level for Latino tracts is comparable with that in White tracts. Our findings suggest that home valuation differences between minority and White tracts do exist in terms of both likelihood and severity, even after controlling for important factors.⁴³ Therefore, it appears that some special effect associated with minority tracts is making a difference in the appraisal process. One idea to address these differences is to test whether automated valuation tools can help mitigate the special effect associated with Black or Latino tracts. If these tools help reduce the differences in terms of likelihood and/or severity, policy makers should consider leveraging them more often as alternatives to traditional appraisals in a safe and sound manner.

Given such a persistent gap between Black (Latino) tracts and White tracts, a natural follow-up question is to investigate the causal mechanisms that drive the gap. In residential appraisals, appraisers develop home values primarily by examining comps, so we have tried checking different perspectives of comps. First, we examine whether comp distance can help explain the gaps. Our initial research indicates that the raw average comp distance is much smaller when the subject property is in a Black or Latino tract than in a White tract, as reported in our first Research Note. However, when using a modeling approach, we find that this difference disappears after controlling for important factors that affect home values and appraisal practices. This finding implies that average comp distance is not likely to be a driving force for different appraisal outcomes. Similarly, we estimate models by

One idea to address these differences is to test whether automated valuation tools can help mitigate the special effect associated with Black or Latino tracts.

-
- 41 When interpreting the findings, one should keep in mind that this research is based purchase appraisals from 2015 to 2000. During this research period, house prices either remained stable or moderately increased, which is different from the rapidly increasing home prices observed in recent markets. Also, purchase and refinance are two different market transactions, and thus the pattern of appraisal outcomes may differ due to the different nature of these transactions.
- 42 It’s critical to control for a list of appropriate variables in this type of modeling research. On the one hand, we should include factors that impact house values or shape appraisal practices, such as number of bedrooms and location effects. On the other hand, we should not over-control by adding factors like credit score because credit score is not supposed to be a factor during the appraisal process. These factors can be called “tainted variables” in employment discrimination case law (Hersch & Bullock, 2014).
- 43 Conceptually, there are possible explanations for the remaining racial and ethnic discrepancies such as omitted variables and model misspecification; however, researchers and practitioners can interpret the consistent results from multiple models in our research by considering that a list of important factors are already controlled for. Ambrose et al. (2021) find that the race of the appraiser and borrower are not related to valuations in a nationwide sample of subprime mortgages originated from 2000 to 2007. They, however, do not examine whether there is systematic appraisal bias based on the racial composition of the neighborhood.



using comp reconciliation, comp variance, and the ratio between the highest (or lowest) comp and contract price as the dependent variable respectively, but do not observe any significant differences between minority tracts and White tracts. More research is needed to better understand how the gaps occur.

Another interesting topic relates to the mixed impacts that different appraisal outcomes have on borrowers. Fout et al. (2021) and Fout and Yao (2016) find that receiving an appraisal lower than the contract price is not necessarily bad for borrowers because it leads to a higher probability of renegotiating to a lower price. On the other hand, Fout and Yao (2016) also find that it increases the probability of the contract being delayed or even cancelled. From a seller's perspective, "appraisal value lower than contract price" is certain to be harmful because either the house will be sold at a lower price, or the current contract will be delayed or cancelled. Fundamentally, different appraisal outcomes may have profound influences on wealth accumulation through homeownership in minority communities. If houses in minority neighborhoods are more likely to be appraised below the contract price, they are more likely to be transacted at a lower price, thus becoming comps with lower prices in future house sales. Due to these perpetuating effects, policy makers need to know how varying appraisal results dynamically affect wealth accumulation through homeownership in minority neighborhoods and enact appropriate interventions if warranted.



References

- Airgood-Obrycki, W., & Rieger, S. (2019). Defining suburbs: How definitions shape the suburban landscape. *Joint Center for Housing Studies, Harvard University*. https://www.jchs.harvard.edu/sites/default/files/Harvard_JCHS_Airgood-Obrycki_Rieger_Defining_Suburbs.pdf.
- Ambrose, B. W., Conklin, J., Coulson, N. E., Diop, M., & Lopez, L. A. (2021). Does Appraiser and Borrower Race Affect Valuation? *Available at SSRN 3951587*.
- Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics*. Princeton University Press.
- Brummet, Q., and D. Reed. 2019. "The Effects of Gentrification on the Well-being and Opportunity of Original Resident Adults and Children." Federal Reserve Board of Philadelphia Working Paper No. 19-30, Available at SSRN: <https://ssrn.com/abstract=3421581> or <http://dx.doi.org/10.21799/frbp.wp.2019.30>.
- Calem, P. S., L. Lambie-Hanson, and L. I. Nakamura. 2017. "Appraising Home Purchase Appraisals." Working Paper No. 17-23, Federal Reserve Bank of Philadelphia.
- Fout, H., N. Mota, and E. Rosenblatt. 2021. "When Appraisers Go Low, Contracts Go Lower: The Impact of Expert Opinions on Transaction Prices." *The Journal of Real Estate Finance and Economics* 1–41. doi: 10.1007/s11146-020-09800-6.
- Fout, H., & Yao, V. (2016). Housing market effects of appraising below contract. Working paper, available at: https://www.researchgate.net/publication/298807852_Housing_Market_Effects_of_Appraising_Below_Contract.
- Gould Ellen, I., & O'Regan, K. (2008). Reversal of fortunes? Lower-income urban neighbourhoods in the US in the 1990s. *Urban Studies*, 45(4), 845-869.
- Hersch, J., Bullock, B.(2014). The Use and Misuse of Econometric Evidence in Employment Discrimination Cases. *Washington and Lee Law Review*, 71(4), Article 7
- Hipsman, Nathaniel E. "Race and Home Price Appreciation in the United States: 1992–2012." (2018).
- Kermani, Amir, and Francis Wong. Racial Disparities in Housing Returns. No. w29306. National Bureau of Economic Research, 2021
- Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics*, 95(2), 391-413.
- Muellbauer, J. (1974). Household Production Theory, Quality, and the " Hedonic Technique". *The American Economic Review*, 64(6), 977-994.
- Neal, M., Stochak, S., Zhu, L., & Young, C. (2020). How Automated Valuation Models Can Disproportionately Affect Majority-Black Neighborhoods. *Urban Institutes working paper*.
- Perry, A., Rothwell, J., & Harshbarger, D. (2018). The devaluation of assets in black neighborhoods: The Case of Residential Property, Brookings Metropolitan Policy Program.
- Pinto, E., & Peter, T. (2021). AEI Housing Center Critique of Freddie Mac's Note on "Racial and Ethnic Valuation Gaps in Home Purchase Appraisals". *AEI working paper*.
- Richardson, J., Mitchell, B., & Franco, J. (2019). Shifting neighborhoods: Gentrification and cultural displacement in American cities. *NCRC working paper*.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55.



Appendix

APPENDIX 1:

Full modeling results for Exhibit 1

	Depvar: appraisal lower than price					
	Black			Latino		
Black tract flag	0.049*** (0.009)	0.042*** (0.008)	0.024*** (0.003)			
Latino tract flag				0.080*** (0.014)	0.062*** (0.014)	0.029*** (0.003)
living area 1,000 square feet		-0.016*** (0.001)	-0.019*** (0.001)		-0.017*** (0.001)	-0.020*** (0.001)
fireplace		-0.004* (0.002)	0.005*** (0.001)		-0.004* (0.002)	0.005*** (0.000)
pool		0.028*** (0.002)	0.013*** (0.001)		0.028*** (0.002)	0.013*** (0.001)
garage		-0.001 (0.002)	0.001 (0.001)		-0.004* (0.002)	0.000 (0.001)
No. baths: 2		0.009*** (0.002)	0.001 (0.001)		0.010*** (0.002)	0.001 (0.001)
No. baths: > 2		0.010*** (0.003)	-0.000 (0.001)		0.011*** (0.002)	-0.001 (0.001)
No. bedrooms: 3		-0.001 (0.001)	-0.004*** (0.001)		-0.001 (0.001)	-0.004*** (0.001)
No. bedrooms: >= 4		-0.002 (0.002)	-0.007*** (0.001)		-0.001 (0.002)	-0.007*** (0.001)
No. stories: > 1		-0.008*** (0.001)	-0.000 (0.001)		-0.008*** (0.001)	-0.000 (0.001)
construction year: 1961-1980		0.012*** (0.002)	0.009*** (0.001)		0.013*** (0.002)	0.009*** (0.001)
construction year: 1981-2000		0.026*** (0.002)	0.020*** (0.001)		0.027*** (0.002)	0.021*** (0.001)
construction year: 2001-2010		0.037*** (0.002)	0.032*** (0.001)		0.038*** (0.002)	0.032*** (0.001)
construction year: 2011-2020		0.007*** (0.002)	0.009*** (0.001)		0.007*** (0.002)	0.009*** (0.001)
urban		0.018*** (0.003)	0.007*** (0.001)		0.016*** (0.003)	0.007*** (0.001)
suburban		0.022*** (0.003)	0.008*** (0.001)		0.022*** (0.002)	0.008*** (0.001)
log median household income		-0.010*** (0.003)	-0.021*** (0.002)		-0.010*** (0.003)	-0.022*** (0.002)
median age		-0.000 (0.000)	-0.000*** (0.000)		0.000 (0.000)	-0.000*** (0.000)
share of household with child(ren)		0.022***	0.002		0.021***	0.006***



		(0.005)	(0.002)		(0.004)	(0.002)
share of population in labor force		-0.018*	0.036***		-0.022*	0.031***
		(0.011)	(0.004)		(0.012)	(0.004)
share of owner-occupied units		-0.017***	0.011***		-0.017***	0.012***
		(0.006)	(0.003)		(0.006)	(0.003)
low FSD		0.004***	-0.000		0.004***	-0.001
		(0.001)	(0.001)		(0.001)	(0.001)
median FSD		-0.000	-0.002***		-0.001	-0.003***
		(0.001)	(0.001)		(0.001)	(0.000)
missing FSD		-0.001	-0.009***		-0.002	-0.011***
		(0.002)	(0.001)		(0.002)	(0.001)
gentrification		-0.005	-0.005***		-0.002	-0.003**
		(0.003)	(0.002)		(0.003)	(0.002)
turn-over rate		0.000**	-0.000***		0.000***	-0.000***
		(0.000)	(0.000)		(0.000)	(0.000)
FHA loan indicator		0.010***	0.002***		0.009***	0.002***
		(0.001)	(0.001)		(0.001)	(0.001)
constant	0.073***	0.175***		0.073***	0.179***	
	(0.002)	(0.030)		(0.002)	(0.028)	
Year and Month Dummies	No	Yes	Yes	No	Yes	Yes
Appraiser-CBSA Fixed Effects	No	No	Yes	No	No	Yes
Depvar Mean of White Tracts	0.073	0.073	0.073	0.073	0.073	0.073
Adjusted R ²	0.001	0.010	0.076	0.004	0.014	0.079
N	10899546	10898846	10698843	11081215	11080594	10878762

*** 0.01 ** 0.05 * 0.1

Note: The dependent variable is whether appraisal is lower than contract price. The minority tract flag is created based on 50% threshold. We've put on the right-hand side property-level controls, including house living area, fireplace, pool, garage, baths, bedrooms, stories, construction year, FSD categories and FHA loan indicator, and tract-level controls, including urban/suburban flags, median household income, median age, share of households with children, share of population in labor force, share of owner-occupied units, gentrification indicator and turn-over rate. Standard errors are clustered at county level. We used binning to transform the following variables into categorical features as binning introduces non-linearity and tends to improve model performance: baths, bedrooms, stories, construction year, median age, share of households with children, share of population in labor force. We have done robustness checks by using the continuous variables and our conclusions hold well.



APPENDIX 2:

Full modeling results for Exhibit 2

	Depvar: appraisal lower than price					
	Black			Latino		
Black tract flag	0.5678***	0.2987***	0.4181***			
Latino tract flag				0.8273***	0.3193***	0.2620***
fire place		0.0811***	0.1192***		0.0803***	0.0767***
pool		0.2143***	0.1446***		0.2090***	0.2107***
garage		0.0222***	0.0374***		0.0128***	0.0241**
gentrification		-0.0714***	-0.1073***		-0.0310***	-0.0151
FHA appraisal indicator		0.0614***	0.0171		0.0607***	0.0365***
living area square feet in 1000		-0.3650***	-0.3477***		-0.3650***	-0.3395***
log median household income		-0.2840***	-0.3338***		-0.2912***	-0.3985***
share of owner occupied units		0.0814***	0.1753***		0.1120***	0.3205***
turn over rate		-0.0000***	-0.0001***		-0.0000***	-0.0000
No. baths: >=2 and <3.5		0.0220***	0.0749***		0.0218***	0.0041
No. baths: >=3.5 and <=4		0.0928***	0.1310***		0.0904***	0.0722***
No. bedrooms: 2-10		-0.2554***	-0.3739***		-0.2452***	-0.2336***
No. stories: >=1.5 and <2		-0.0476***	-0.1309***		-0.0500***	-0.1872***
No. stories: >=2 and <=4		0.0190***	0.0031		0.0196***	0.0189***
construction year: 1948-1983		0.1235***	0.1836***		0.1188***	0.1364***
construction year: 1984-1991		0.2774***	0.4114***		0.2761***	0.3055***
construction year: 1992-2007		0.4398***	0.5777***		0.4349***	0.4144***
construction year: 2008-2020		0.1818***	0.2461***		0.1726***	0.1054***
suburban		0.0296***	0.0111		0.0309***	0.0401***
rural		-0.1527***	-0.2180***		-0.1332***	-0.0565***
median age 30-47.1		-0.0202***	-0.0283		-0.0320***	-0.0724***
median age >47.1		-0.0843***	-0.1058***		-0.0971***	-0.1900***
share of household with child(ren) 0.3-0.8		0.0693***	0.0647***		0.0740***	0.1109***
share of household with child(ren) >=0.8		0.0900***	0.0338*		0.1088***	0.1461***
share of labor forces: 0.4-0.8		0.1671***	0.1224***		0.1628***	0.3075***
share of labor forces: >=0.8		0.1655***	0.1434***		0.1387***	0.2229***
missing FSD		-0.1867***	-0.1162***		-0.2040***	-0.1074***
low FSD		0.0434***	0.0489***		0.0351***	0.1331***
median FSD		-0.0156***	-0.0001		-0.0272***	0.0807***
intercept	-2.5380***	0.5662***	2.5502***	-2.5380***	-2.5525***	-1.2764***
Year and Month Dummies	No	Yes	Yes	No	Yes	Yes
CBSA Fixed Effects	No	Yes	Yes	No	Yes	Yes
Appraiser Fixed Effects	No	No	Yes	No	No	Yes
C-Statistic	0.511	0.658	0.724	0.526	0.664	0.712

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: The dependent variable is whether appraisal is lower than contract price. The minority tract flag is created based on 50% threshold. We've put on the right-hand side property-level controls, including house living area, fireplace, pool, garage, baths, bedrooms, stories, construction year, FSD categories and FHA loan indicator, and tract-level controls, including urban/suburban flags, median household income, median age, share of households with children, share of population in labor force, share of owner-occupied units, gentrification indicator and turn-over rate. We used binning to transform the following variables into categorical features as binning introduces non-linearity and tends to improve model performance: baths, bedrooms, stories, construction year, median age, share of households with children, share of population in labor force. We have done robustness checks by using the continuous variables and our conclusions hold well.



APPENDIX 3:

Full modeling results for Exhibit 4

	Depvar: appraisal lower than price		
Black Share (%)	0.0008*** (0.0001)	0.0009*** (0.0001)	0.0005*** (0.0001)
Latino Share (%)	0.0015*** (0.0002)	0.0014*** (0.0002)	0.0006*** (0.0000)
living area 1,000 square feet		-0.0183*** (0.0009)	-0.0207*** (0.0008)
fireplace		-0.0018 (0.0024)	0.0059*** (0.0005)
pool		0.0257*** (0.0022)	0.0128*** (0.0015)
garage		-0.0001 (0.0024)	0.0006 (0.0007)
No. baths: 2		0.0080*** (0.0022)	0.0012** (0.0006)
No. baths: > 2		0.0095*** (0.0027)	0.0001 (0.0008)
No. bedrooms: 3		-0.0023 (0.0015)	-0.0041*** (0.0006)
No. bedrooms: >= 4		-0.0030* (0.0018)	-0.0066*** (0.0008)
No. stories: > 1		-0.0064*** (0.0013)	-0.0002 (0.0010)
construction year: 1961-1980		0.0141*** (0.0017)	0.0110*** (0.0010)
construction year: 1981-2000		0.0289*** (0.0025)	0.0232*** (0.0014)
construction year: 2001-2010		0.0394*** (0.0028)	0.0346*** (0.0014)
construction year: 2011-2020		0.0037 (0.0025)	0.0077*** (0.0013)
urban		0.0102*** (0.0028)	0.0056*** (0.0015)
suburban		0.0197*** (0.0024)	0.0085*** (0.0014)
log median household income		-0.0041 (0.0031)	-0.0171*** (0.0022)
median age		0.0004* (0.0002)	-0.0002*** (0.0001)
share of household with child(ren)		0.0145*** (0.0035)	0.0034 (0.0025)
share of population in labor force		-0.0019 (0.0131)	0.0289*** (0.0042)
share of owner-occupied units		-0.0046	0.0151***



		(0.0060)	(0.0029)
low FSD		0.0034***	0.0001
		(0.0011)	(0.0007)
median FSD		-0.0010	-0.0019***
		(0.0007)	(0.0006)
missing FSD		-0.0053***	-0.0113***
		(0.0019)	(0.0011)
gentrification		-0.0130***	-0.0072***
		(0.0026)	(0.0015)
turn-over rate		0.0000	-0.0000***
		(0.0000)	(0.0000)
FHA loan indicator		0.0085***	0.0029***
		(0.0012)	(0.0007)
constant	0.0560***	0.0650**	
	(0.0021)	(0.0302)	
Year and Month Dummies	No	Yes	Yes
Appraiser-CBSA Fixed Effects	No	No	Yes
Depvar Mean	0.083	0.083	0.083
Adjusted R ²	0.010	0.019	0.083
N	12623454	12622663	12406292

*** 0.01 ** 0.05 * 0.1

Note: The dependent variable is whether appraisal is lower than contract price. The treatment variable is minority share. We've put on the right-hand side property-level controls, including house living area, fireplace, pool, garage, baths, bedrooms, stories, construction year, FSD categories and FHA loan indicator, and tract-level controls, including urban/suburban flags, median household income, median age, share of households with children, share of population in labor force, share of owner-occupied units, gentrification indicator and turn-over rate. Standard errors are clustered at county level. We used binning to transform the following variables into categorical features as binning introduces non-linearity and tends to improve model performance: baths, bedrooms, stories, construction year, median age, share of households with children, share of population in labor force. We have done robustness checks by using the continuous variables and our conclusions hold well.



APPENDIX 4:

Full modeling results for Exhibit 5

	Depvar: appraisal lower than price		
Black share	0.0098***	0.0064***	0.0076***
Latino share	0.0160***	0.0077***	0.0070***
fire place		0.0909***	0.1061***
pool		0.1981***	0.1902***
garage		0.0278***	0.0381***
gentrification		-0.0935***	-0.0646***
FHA appraisal indicator		0.0666***	0.0361***
living area square feet in 1000		-0.3602***	-0.3455***
log median household income		-0.1964***	-0.2798***
share of owner occupied units		0.1680***	0.2673***
turn over rate		-0.0000***	-0.0000***
No. baths: >=2 and <3.5		0.0271***	0.0501***
No. baths: >=3.5 and <=7		0.1034***	0.1226***
No. bedrooms: 2-10		-0.2544***	-0.2970***
No. stories: >=1.5 and <2		-0.0469***	-0.1306***
No. stories: >=2 and <=4		0.0261***	0.0254***
construction year: 1948-1983		0.1466***	0.1838***
construction year: 1984-1991		0.3037***	0.3575***
construction year: 1992-2007		0.4624***	0.5018***
construction year: 2008-2020		0.1605***	0.1504***
suburban		0.0505***	0.0418***
rural		-0.1084***	-0.0695***
median age 30-47.1		-0.0207***	-0.0390***
median age >47.1		-0.0621***	-0.0960***
share of household with child(ren) 0.3-0.8		0.0418***	0.0716***
share of household with child(ren) >=0.8		0.0447***	0.0678***
share of labor forces: 0.4-0.8		0.1302***	0.2501***
share of labor forces: >=0.8		0.1154***	0.2256***
missing FSD		-0.2042***	-0.1450***
low FSD		0.0512***	0.0999***
median FSD		-0.0124***	0.0296**
Intercept	-2.7388***	-2.5571***	1.7873***
Year and Month Dummies	No	Yes	Yes
CBSA Fixed Effects	No	Yes	Yes
Appraiser Fixed Effects	No	No	Yes
C-Statistic	0.595	0.67	0.716

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: The dependent variable is whether appraisal is lower than contract price. The treatment variable is minority share. We've put on the right-hand side property-level controls, including house living area, fireplace, pool, garage, baths, bedrooms, stories, construction year, FSD categories and FHA loan indicator, and tract-level controls, including urban/suburban flags, median household income, median age, share of households with children, share of population in labor force, share of owner-occupied units, gentrification indicator and turn-over rate. We used binning to transform the following variables into categorical features as binning introduces non-linearity and tends to improve model performance: baths, bedrooms, stories, construction year, median age, share of households with children, share of population in labor force. We have done robustness checks by using the continuous variables and our conclusions hold well.



APPENDIX 5:

Full modeling results for Exhibit 7

	Depvar: (appraisal – price)/price x 100%					
	Black			Latino		
Black tract flag	-1.435*** (0.137)	-0.900*** (0.082)	-0.741*** (0.052)			
Latino tract flag				-0.319*** (0.103)	-0.179* (0.096)	-0.183*** (0.062)
living area 1,000 square feet		-0.437*** (0.020)	-0.337*** (0.016)		-0.446*** (0.020)	-0.334*** (0.016)
fireplace		0.008 (0.022)	-0.009 (0.014)		0.031 (0.024)	-0.004 (0.014)
pool		-0.339*** (0.032)	-0.294*** (0.024)		-0.346*** (0.032)	-0.303*** (0.023)
garage		0.307*** (0.044)	0.426*** (0.026)		0.364*** (0.041)	0.424*** (0.028)
No. baths: 2		0.236*** (0.037)	0.220*** (0.026)		0.214*** (0.033)	0.201*** (0.025)
No. baths: > 2		0.258*** (0.036)	0.247*** (0.025)		0.226*** (0.031)	0.216*** (0.024)
No. bedrooms: 3		0.447*** (0.028)	0.497*** (0.025)		0.462*** (0.026)	0.514*** (0.024)
No. bedrooms: >= 4		0.682*** (0.034)	0.744*** (0.031)		0.688*** (0.032)	0.748*** (0.029)
No. stories: > 1		0.118*** (0.024)	0.128*** (0.018)		0.107*** (0.021)	0.119*** (0.017)
construction year: 1961-1980		0.500*** (0.046)	0.406*** (0.035)		0.457*** (0.044)	0.389*** (0.030)
construction year: 1981-2000		0.840*** (0.053)	0.731*** (0.033)		0.794*** (0.052)	0.716*** (0.032)
construction year: 2001-2010		1.124*** (0.057)	0.996*** (0.035)		1.063*** (0.056)	0.973*** (0.035)
construction year: 2011-2020		1.741*** (0.059)	1.464*** (0.040)		1.680*** (0.056)	1.433*** (0.039)
urban		0.171*** (0.057)	0.305*** (0.044)		0.165*** (0.055)	0.328*** (0.045)
suburban		0.053 (0.052)	0.269*** (0.041)		0.055 (0.048)	0.279*** (0.041)
log median household income		0.880*** (0.071)	0.728*** (0.044)		0.922*** (0.066)	0.754*** (0.043)
median age		-0.006** (0.003)	-0.001 (0.002)		-0.008*** (0.003)	-0.001 (0.002)
share of household with child(ren)		-0.042 (0.085)	-0.066 (0.049)		-0.063 (0.082)	-0.091 (0.055)
share of population in labor force		0.967*** (0.195)	0.481*** (0.124)		0.780*** (0.194)	0.432*** (0.120)



share of owner-occupied units	-0.554***	-0.312***		-0.571***	-0.321***	
	(0.128)	(0.076)		(0.121)	(0.074)	
low FSD	1.196***	1.005***		1.216***	1.031***	
	(0.039)	(0.034)		(0.036)	(0.034)	
median FSD	0.425***	0.370***		0.420***	0.369***	
	(0.035)	(0.036)		(0.036)	(0.035)	
missing FSD	0.849***	0.746***		0.870***	0.797***	
	(0.041)	(0.036)		(0.043)	(0.040)	
gentrification	-0.163**	-0.153***		-0.136*	-0.107*	
	(0.064)	(0.053)		(0.074)	(0.059)	
turn-over rate	0.001***	0.001***		0.001***	0.001***	
	(0.000)	(0.000)		(0.000)	(0.000)	
FHA loan indicator	0.090***	0.135***		0.099***	0.123***	
	(0.023)	(0.017)		(0.022)	(0.016)	
constant	-4.431***	-16.690***		-4.431***	-16.951***	
	(0.044)	(0.736)		(0.044)	(0.671)	
Year and Month Dummies	No	Yes	Yes	No	Yes	Yes
Appraiser-CBSA Fixed Effects	No	No	Yes	No	No	Yes
Depvar Mean of White Tracts	-4.431	-4.431	-4.431	-4.431	-4.431	-4.431
Adjusted R ²	0.008	0.071	0.186	0.001	0.064	0.179
N	816338	816289	790162	855596	855556	828911

*** 0.01 ** 0.05 * 0.1

Note: The dependent variable is the percentage difference between appraisal and contract price. The minority tract flag is created based on 50% threshold. We've put on the right-hand side property-level controls, including house living area, fireplace, pool, garage, baths, bedrooms, stories, construction year, FSD categories and FHA loan indicator, and tract-level controls, including urban/suburban flags, median household income, median age, share of households with children, share of population in labor force, share of owner-occupied units, gentrification indicator and turn-over rate. Standard errors are clustered at county level. We used binning to transform the following variables into categorical features as binning introduces non-linearity and tends to improve model performance: baths, bedrooms, stories, construction year, median age, share of households with children, share of population in labor force. We have done robustness checks by using the continuous variables and our conclusions hold well.



Prepared by Modeling, Econometrics, Data Science & Analytics, and Single-Family Risk Management

Melissa Narragon, Senior Director of Fair Lending Analytics, Modeling, Econometrics, Data Science & Analytics

Danny Wiley, Senior Director of Single Family Property Valuation, Single Family Risk Management

Vivian Li, Director of Fair Lending Analytics, Modeling, Econometrics, Data Science & Analytics

Zhiqiang Bi, Director of Collateral Modeling, Modeling, Econometrics, Data Science & Analytics

Kangli Li, Quantitative Analytics Senior, Modeling, Econometrics, Data Science & Analytics

Xue Wu, Quantitative Analytics Senior, Modeling, Econometrics, Data Science & Analytics

www.freddiemac.com/research

Opinions, estimates, forecasts, and other views contained in this document are those of Freddie Mac's economists and other researchers, do not necessarily represent the views of Freddie Mac or its management, and should not be construed as indicating Freddie Mac's business prospects or expected results. Although the authors attempt to provide reliable, useful information, they do not guarantee that the information or other content in this document is accurate, current or suitable for any particular purpose. All content is subject to change without notice. All content is provided on an "as is" basis, with no warranties of any kind whatsoever. Information from this document may be used with proper attribution. Alteration of this document or its content is strictly prohibited.

© 2022 by Freddie Mac