Heliyon 12 (2026) e45520

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Philadelphia then and now, disparities in mortgage lending and sensitivity to unmeasured confounders☆

Maia Berkane *, Diego Alvarez, Nico Tanzi [1]

Research and Analytics, Kevin D. Oden & Associates, San Francisco, USA

ARTICLE INFO

Keywords:
Mortgage lending
Loan denial
Racial disparities
Unmeasured confounders
Sensitivity analysis
E-Value
HMDA
Philadelphia MSA
Logistic regression
Marginal effects

ABSTRACT

Racial and ethnic disparities in mortgage lending have been documented across the United States for decades, yet their persistence and causal robustness remain subjects of ongoing debate. A central challenge in this literature is that publicly available lending data — including the Home Mortgage Disclosure Act (HMDA) data — lacks key underwriting variables such as credit scores, making it difficult to distinguish discriminatory treatment from differences in creditworthiness. This paper examines trends in mortgage denial disparities between minority and White applicants in the Philadelphia Metropolitan Statistical Area (MSA) from 2017 to 2022, and introduces the E-value sensitivity framework — widely adopted in epidemiology but not previously applied to fair lending analysis — to quantify the strength of unmeasured confounding required to explain away observed disparities.

Black applicants face denial odds approximately 2.5 times those of similarly situated White applicants each year, with a consistent marginal effect of approximately seven percentage points. E-values for Black applicants remain above 4.0 based on the odds ratio and above 3.0 based on marginal effects throughout the study period, indicating that an unmeasured confounder would need to be associated with both race and denial by at least a fourfold relative risk — above and beyond measured covariates — to account for the observed disparity. We demonstrate that FICO score, the most plausible omitted confounder, is unlikely to satisfy this threshold. Disparities for Hispanic and Asian applicants, while significant in most years, show lower E-values and greater susceptibility to unmeasured confounding. These findings suggest that the persistent Black-White denial gap in Philadelphia is unlikely to be fully explained by unobserved risk factors and warrants continued regulatory attention.

1. Introduction

Home ownership is a way to build wealth overtime through capital gain, tax savings and price appreciation. However, minority groups, especially non-Hispanic African Americans, were overtly discriminated against and denied mortgage loans. On October 28,

☆ The work described has not been published previously except in the form of a preprint. A summary of this research was posted on the authors' organizational website (kdoden.com/mortgage-lending) in May 2024, after the manuscript was first submitted to Heliyon on April 4, 2024. The website is owned by Kevin D. Oden & Associates, the authors' own organization.

* Corresponding author.
E-mail address: maia.berkane@kdoden.com (M. Berkane).
[1] Current affiliation: Muttdata, Argentina.

https://doi.org/10.1016/j.heliyon.2026.e45520
Received 14 April 2024; Received in revised form 17 September 2026; Accepted 25 September 2026
Available online 3 October 2026
2405-8440/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1974, the United States Congress enacted the Equal Credit Opportunity Act (ECOA) codified at 15 U S C. §1691 et seq [1]. It is a law that makes it illegal for any creditor to discriminate against any applicant, with respect to any aspect of a credit transaction, on the basis of race, color, religion, national origin, sex, marital status, or age (provided the applicant has the capacity to repay the loan). Quillian et al. [2] used a meta-analysis to investigate trends in mortgage underwriting and pricing discrimination against African Americans, Hispanic and Asian, from 1976 to 2016. They showed that while the disparities have declined slightly since the implementation of ECOA, they remained high and persistent. It is interesting to investigate if the trend of disparities has persisted since 2017. In 2020, the Federal Reserve Bank of Philadelphia stated that, although the African American-White homeownership gap was smaller than the U.S. national average, there has been little progress to close that gap over the last three decades [3]. From 2017 to 2021, ESSA Bank & Trust was engaging in discriminatory lending in the Philadelphia MSA and agreed to pay over $3 million to settle a federal redlining lawsuit brought by the Department of Justice in 2023 [4]. In July 2022, the Consumer Financial Protection Bureau (CFPB) ordered Trident to pay more than $22 million for deliberate discrimination against minority families and violation of ECOA [5]. It is reasonable to ask whether the Philadelphia MSA differs from the other MSAs in terms of the disparities in mortgage lending between White and minority applicants. We address this question in a companion study currently under review, Berkane [6]. Philadelphia MSA is a very large MSA with 83,178 mortgage applications in 2016, $15,632\ (18.8\%)$ of which were Black, $4,729\ (5.7\%)$ Asian, $4,614\ (5.5\%)$ Hispanic, $37,868\ (45.5\%)$ White, 19,858 (23.9%) with missing race information, and the remainder belonging to other race or ethnicity categories [7]. Since the HMDA database does not include crucial attributes for lending decision making, such as the Fair Isaac Corporation (FICO) score, we treat these attributes as unmeasured confounders and perform sensitivity analysis to assess the strength of the race and ethnicity effect on the lending decision. We analyze the Philadelphia MSA HMDA data from 2017 to 2022, which results in a dataset large enough to allow reasonable power for our statistical tests. To our knowledge, none of the published articles on mortgage underwriting or pricing discrimination have provided a sensitivity analysis to unmeasured confounders, nor have we seen some of the advances in other fields faced with observational data, for instance Epidemiology, applied to HMDA data analysis. The evidence of causality, termed E-value, of Ding and VanderWeele [8] is now widely adopted in Epidemiological research and published in addition to the p-value. In this paper, we ask the following questions:

1. What is the trend of the disparities in conventional, first lien, 1-4 family homes, mortgage lending decision between minority and White applicants?
2. What is the minimum strength an unmeasured confounder needs to satisfy with both minority groups and the outcome, to fully explain away the disparities in mortgage underwriting year after year?

To answer these questions, each year, we perform the following:

1. Data processing, transformations, filtering and variables creation.
2. A logistic regression of the denial of loans, on minority group indicators as well as the missing race indicators, and adjust for all the relevant and available variables in HMDA data.
3. A sensitivity analysis to unmeasured confounders using the method of Ding and VanderWeele [8], namely the E-value.

Section 2 reviews the literature, section 3 gives details of the data, its processing, transformation and variables creation, section 4 provides details of the methodologies used in the paper, section 5 gives the results for each year, the trend of disparities, as well as sensitivity to unmeasured confounders. Section 6 discusses the results; section 7 presents the limitations. We conclude the paper, in section 8, with final thoughts on our analysis.

2. Literature review

Research has shown that racial-based disparities in the housing markets can take several forms such as disparities in home ownership [9–11], affordable rental housing [9], mortgage processing time [12], and mortgage denial rates [13,14], among other studies. Research has also focused on the methods to investigate these disparities, e.g., the roles of co-applicants [14], and the implications of relative income [13], among other work. Kuebler and Rugh [10] provide a comprehensive analysis of home ownership data in the United States from 2001 to 2010, revealing significant disparities among different racial and ethnic groups. Brooks [9] concentrated on the disparities in affordable rental housing among different ethno-racial groups in the United States. His findings revealed that White households have higher rates of affordable housing than Black and Hispanic households. These disparities, despite housing policies, remained nearly identical between 2005 and 2019. Lending disparities are also reflected in mortgage processing time, as Wei and Zhao [12] examined. Their study finds that Black borrowers are underrepresented (i.e., the percentage of Black applicants that apply for a mortgage loan is lower than the percentage of White ones) and experience a longer processing time than White borrowers for the mortgages securitized by government-sponsored enterprises (GSE).

Loya [13,14] further underscores the complexity of analyzing the lending disparities. He highlights the importance of some measured but sometimes ignored confounders such as the role of co-applicants and the role of relative income. Based on HMDA data, his findings reveal that interracial couples involving a Black or Latino co-applicant are more likely to experience a high-cost loan or be denied a mortgage, than mono-racial White couples. Additionally, he found that high-cost loan origination and loan rejection are higher among Latino and Black applicants across relative income levels compared to White applicants. This introduces another layer of complexity, since he suggests that the disparities in loan outcomes are not only a function of individual characteristics but also of the dynamics within couples and the relation between the individual characteristics and their neighborhood (e.g., relative income).

Namin et al. [15] explore the persistence of mortgage lending bias in the United States, 80 years after the creation of the Homeowner's Loan Corporation (HOLC) "residential security maps." HOLC grades were used to assess the creditworthiness and investment risk of neighborhoods and showed that the best HOLC grades were often assigned to predominantly white neighborhoods, and worse grades to neighborhoods with high minority concentration. The study compares historic HOLC grades and contemporary levels of mortgage lending bias using spatially detailed HMDA data. Faber [16] makes a directly related argument that racial segregation shapes perceived creditworthiness and creates geographic inequality in mortgage access in the post-recession period. Hammel et al. [17] examines how high foreclosure rates in neighborhoods, which are themselves racially concentrated, increase the probability of denial for minority applicants in the post-recession period. These findings collectively underscore the geographic dimension of lending disparities, which motivates our focus on the Philadelphia MSA as a specific and historically documented case.

Haupert [18,19] draws on HMDA data across the 200 largest U.S. metropolitan areas to examine how racial and ethnic residential segregation shapes subprime lending outcomes. His findings indicate that both traditional and fintech lenders originate more subprime loans in areas with greater concentrations of Black and Hispanic residents. Fintech lenders, however, show a more complex pattern: they issue fewer subprime loans than traditional lenders in highly segregated Black communities, while the relationship reverses in areas where White and Hispanic populations are more residentially separated.

As an observational data, the HMDA data comes with problems such as selection bias (people chose to apply for mortgages and are not a random sample of the Philadelphia MSA), measurement errors and unmeasured confounders. HMDA data is a common source of information for researchers to investigate the behavior of mortgage underwriting and the progress in reduction of racial and ethnic disparities in lending and pricing decisions. However, that data, at least the non-confidential, publicly available one, lacks crucial credit worthiness variables that lenders rely on to make their decisions. FICO score, an important credit variable, is not available in HMDA data and is known to be correlated with the minority groups as well as the mortgage underwriting decision. In 2021, Forbes [20] reported that more than 1 in 5 Black consumers and 1 in 9 Hispanic consumers had FICO scores below 620; meanwhile, 1 out of every 19 White people are in the sub-620 category".

The Urban Institute [21] in 2023, compared median credit scores in majority Black, majority White, majority Hispanic, and majority Native American communities. The White and Hispanic communities had the highest median scores, while the Black and Native American communities had the lowest. Their table breaks down the credit scores and their ranges for Vantage Score (the basis for the study's scores) and FICO and shows that the median credit score for African Americans is 627 while it is 727 for White Americans and 667 for Hispanic Americans. 627 is considered "near prime" by Vantage Score and 667 and 727 are considered "prime". FICO considers 627 and 667 as "fair" while 727 as "good". Munnell and McEneaney [22] included credit history and other underwriting factors sourced from credit reports and lenders' worksheets and found that Black and Hispanic mortgage applicants in the Boston metropolitan area are still roughly 60 percent more likely to be turned down than White applicants. Ross and Yinger [23] provided an in-depth review of the Munnell and McEneaney [22] Boston Fed Study along with new evidence that the minority-white loan-approval disparities in the Boston data represent discrimination, rather than variation in underwriting standards that can be justified by lenders.

Bhutta et al. [24] attempted to account for unavailable covariates in the HMDA data by incorporating underwriting decision from automatic underwriting systems (AUS), obtained from the lenders, because of their affiliation with the Federal Reserve, and found a residual denial gap of 2% for Black and 1% for Hispanic compared to White applicants. They attributed the racial and ethnic differences in observable risk characteristics to discriminatory treatment elsewhere (e.g., the labor market) or in the inter-generational persistence of historical unequal treatment. They also attributed the difference between their results and those of Munnell and McEneaney [22], to compliance with ECOA by lenders.

Bartlett et al. [25], found a marginal effect of $7-10\%$ but they did not adjust for credit scores and other credit variables. Giacoletti et al. [26] also found a 7% Black-White denial gap without conditioning on credit score, loan-to-value (LTV), or debt-to-income (DTI). Authors such as Munnell and McEneaney [22] and Bhutta et al. [24], were able to obtain confidential credit information of the borrowers, due to their affiliation with the Federal Reserve, and adjusted for FICO scores and other confidential variables. They found the marginal effect for Black applicants was still significant. This confirms the strength of the racial and ethnic effects on underwriting decisions.

In our analysis, we do not adjust for FICO because we cannot obtain that data. However, we do not assume that FICO is the only unmeasured confounder, rather, we estimate the strength of a confounder needed to explain away the minority group effect on the outcome. Sensitivity analysis to unmeasured confounders has been introduced and quantified by Rosenbaum and Rubin [27] under the assumptions that the outcome and the missing confounder are binary variables. Lin et al. [28] extended Rosenbaum and Rubin [27] to non-binary unmeasured confounders but assumed the measured and unmeasured confounders are independent within each exposure group. Ding and VanderWeele [8] derived bounding factors for the true treatment effect that allows assessment of the strength of this effect, without imposing any assumption on the nature of the unmeasured confounder or the outcome.

3. Data

The HMDA data was enacted by the United States Congress in 1975 and implemented by Regulation C. It requires collection, report, and disclosure of information about mortgage lending activities. The information from HMDA is used, among other purposes, for fair lending examination processes, Community Reinvestment Act (CRA) evaluation, and house trends evaluations. The data set consists of a wide range of information from the lending institutions that receive the applications, including the race, ethnicity, and income of the applicant. It also contains the type and size of the loan that is being sought, as well as information about the neighborhood in which the property is located. The data is rigorously scrubbed and validated by the banks before submission, and it must follow specific guidelines from the CFPB. The data used in this study is publicly available and it can be downloaded from the Federal Financial

Institutions Examination Council (FFIEC) Web Site [7]. A description of the data is available on the same site [29]. The HMDA fields and the CFPB standardized fields are provided in Appendix 1.

In this analysis, we restrict the data to applications for conventional, first lien, 1 to 4 family units, for the purpose of living in the house, because it is the most common loan application profile. We also filter the loan decision (action_taken) to the values of 1 (approved) or 3 (denied) and the MSA (msa_md) to code 37964, which is the Philadelphia MSA. We will research other profiles, such as refinancing or home improvement in future work. The total number of applications before and after removing the records with missing covariates used in our analysis, for each year, is given in Table 1 and Fig. 1 shows the racial distribution before and after removing data with those missing covariates.

There doesn't appear to be any impact from the missing data on the relative racial composition of the two data sets.

3.1. Variables transformation and creation

The variables available in the publicly available HMDA data that are relevant to the mortgage underwriting decision and used by several authors mentioned in the literature review section, include loan amount, loan-to-income (LTI), loan-to-value (LTV), the percentage of minority population in the tract, the ratio of the tract income to that of the MSA, the median age of the properties in the tract, the number of owner occupied units in the tract and the number of one to four family homes in the tract. We decided not to use the pricing-related variables because lenders may go through several iterative processes before setting a price. The price the applicant was seeking may not be the one offered by the lender, it is difficult to adjust for that.

We created a dummy variable "denial" equal to 1 when the application is denied (HMDA variable action_taken is equal to 3), "Hispanic", a dummy variable equal to 1 when the HMDA single aggregated ethnicity categorization derived from applicant/borrower and co-applicant/co-borrower ethnicity field (derived_ethnicity) variable is equal to "Hispanic or Latino". When the HMDA single aggregated ethnicity categorization variable derived_ethnicity is not equal to “Hispanic or Latino”, we created the dummy variables "Black", "Asian", "Native", "PacificI" (pacific islander), each equal to 1 when the HMDA single aggregated race categorization derived from applicant/borrower and co-applicant/co-borrower race fields (derived_race) variable is equal to "Black or African American", "Asian", "American Indian or Alaska Native", "Native Hawaiian or Other Pacific Islander", respectively. "RaceNA" is a dummy variable equal to 1 when all the other dummy race variables are equal to 0 and the HMDA single aggregated race categorization variable (derived_race) is not equal to “White.” We also created the dummy variable "no_coapplicant" equal to 1 when the HMDA ethnicity of the co-applicant variable (co.applicant_ethnicity.1) is missing or greater than 2 and the HMDA race of the co-applicant variable (co.applicant_race.1) is missing or greater than 5.

We transformed the HMDA variable “income” to the logarithm of income (log(income)) for 2017 data, and log (income x 1000) for later years because of the way it was reported in HMDA. We also transformed the loan amount variable to its logarithm and labeled it "ln_LoanAmount." We derived LTI as the ratio of loan amount to the income ((loan_amount)/income) for 2017 data, and the ratio of the loan amount to the income, times 1000 (loan_amount/(income x 1000)) for later data. The variable "LTV" is derived as the ratio of the loan amount to the property value (loan_amount)/property_value). Since the property value is only available in the data starting 2018, we constructed it and used it for 2018 and later years. We included the percentage of minority population in the tract (tract_minority_population_percent), the ratio of the tract income to that of the MSA (tract_to_msa_income_percentage) and the number of owner-occupied units in the tract (tract_owner_occupied_units).

In addition to the filters mentioned earlier, observations below the first and above the 99^{th} quantile of income, loan_amount and property_value were replaced by the first and 99^{th} quantile respectively to avoid bias introduced by very large ouliers in those variables. Native or PacificI were excluded because of the minimis size of the sample.

4. Methods

We compared the average of the selected HMDA variables, between Black and White applicants. Their distributions shown in Fig. 2, are different judging by the p-value stemming from *t*-test comparing means of variables for minority versus White applicants. It is thus necessary to remove this imbalance by adjusting through logistic regression, in this case, since the outcome is binary.

Table 1
Number of applications by year before and after removing missing data.

Year	Total Number of Applications	Number of Applications after Removing Missing Covariates[a]
2017	12,916	12,735
2018	14,615	14,171
2019	14,783	14,419
2020	15,566	15,314
2021	18,978	18,569
2022	15,950	15,669

[a] Most of the missing values come from loan-to-value.

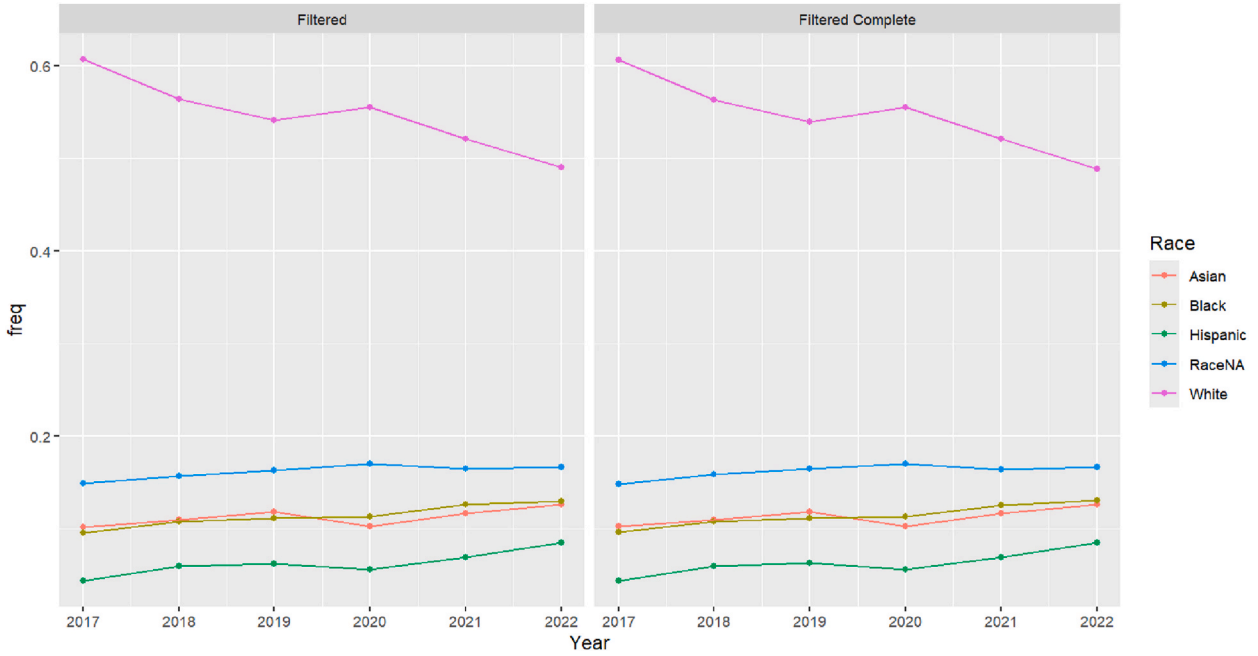


Fig. 1. Distribution of race by year before and after removing missing data.

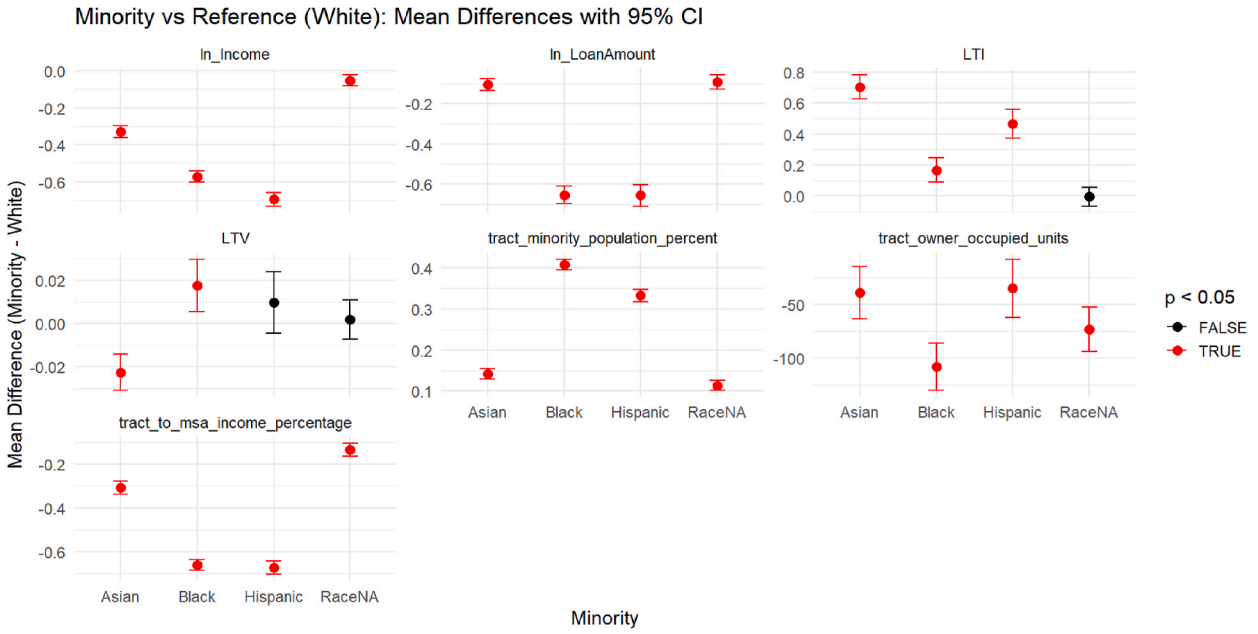


Fig. 2. Imbalance of variables.

4.1. *Logistic regression*

We use logistic regression to estimate the race effect after adjusting for the variables listed above. Given a probability of success p_i for observation Y_i, and predictors vector X_i, logistic regression fits a linear regression to the logit of $p_i, \log(p_i/(1-p_i))$, on X_i. The parameters of the model are estimated by maximizing the log of the likelihood function as follows:

$$L(\beta, Y, X) = \prod_{i=1}^{n} p_i^{Y_i}(1-p_i)^{(1-Y_i)} \tag{1}$$

where

$$\log\left(p_i/(1-p_i)\right) = X_i\beta \tag{2}$$

and

$$p_i = \frac{\exp\left(X_i\beta\right)}{1+\exp\left(X_i\beta\right)} \tag{3}$$

4.2. Sensitivity to unmeasured confounders

Logistic regression allows adjustment of measured confounders only and the exposure effect on the outcome may be biased due to uncontrolled confounders. Sensitivity to unmeasured confounders aims at quantifying the strength of the confounders that will eliminate the exposure effect on the outcome. Ding and VanderWeele [8] provided a minimal threshold of relative risk between the exposure and the confounder as well as between the confounder and the outcome, for an unmeasured confounder to explain away the exposure effect on the outcome. This threshold is termed E-value and has been popular in observational studies, especially in epidemiological studies. We give a brief overview of the E-value and use it to evaluate the effect of unmeasured confounders on the results of the logistic regression.

4.2.1. E-value

The following exposition adopts the notation and framework of Ding and VanderWeele [8] directly, to which the reader is referred for full derivations.

The E-value is a measure related to the "evidence for causality" in observational studies, when they are potentially subject to confounding [8]. As defined by Ding and VanderWeele [8]: "The E-value is the minimum strength of association on the relative risk scale that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured covariates." We will use the same notation as in Ding and VanderWeele [8]. Let E denote the exposure, D denote a binary outcome, C denote the measured confounders, and U denote one or more unmeasured confounders. We will assume the exposure (here, minority) E is binary, but the authors show the results are also applicable to a categorical or continuous exposure and could be applied comparing any two levels of E. Assume that the unmeasured confounder U is categorical with levels $0,1,\ldots,K-1$. However, all the conclusions hold for U of general type. Let the observed relative risk (RR) of the exposure on the outcome given a level c of the measured covariate be

$$RR^{obs}_{ED/c} = \frac{P(D=1|E=1,C=c)}{P(D=1|E=0,C=c)} \tag{4}$$

Let the relative risk of the confounder U at the stratum k, on the exposure E given a level c of the measured covariate be

$$RR_{EU,k/c} = \frac{P(U=k|E=1,C=c)}{P(U=k|E=0,C=c)} \tag{5}$$

Let $RR_{EU|c} = \max_k RR_{EU,k|c}$ be the maximum risk over the stratums of U and for a level c of C. Define

$$RR_{UD|E=1,C=c} = \frac{max_k P(D=1|E=1,U=k,C=c)}{min_k P(D=1|E=1,U=k,C=c)} \tag{6}$$

to be the maximum effect of U on D among the exposed $(E=1)$. Similarly, let

$$RR_{UD|E=0,C=c} = \frac{max_k P(D=1|E=0,U=k,C=c)}{min_k P(D=1|E=0,U=k,C=c)} \tag{7}$$

be the maximum effect of U on D among the unexposed $(E=0)$. Let

$$RR_{UD} = \max\left(RR_{UD|E=1,C=c}, RR_{UD|E=0,C=c}\right) \tag{8}$$

be the maximum effect among the exposed and unexposed and we have omitted the conditioning on the level c of C in RR_{UD} for ease of notation. We also use RR_{EU} instead of $RR_{EU|c}$. The true causal relative risk of the exposure E on the outcome D within stratum c of C is given by

$$RR^{\text{true}}_{ED|c} = \frac{\Sigma_k P(D=1|E=1,C=c,U=k)P(U=k|C=c)}{\Sigma_k P((D=1|E=0,C=c,U=k)P(U=k|C=c))} \tag{9}$$

Ding and VanderWeele [8] show that $RR^{\text{true}}_{ED|c}$ is bounded below by

$$RR_{ED|c}^{\text{true}} \geq RR_{ED|c}^{obs} \frac{1}{\frac{RR_{EU} \times RR_{UD}}{RR_{EU} + RR_{UD} - 1}} \tag{10}$$

The bounding factor $\frac{RR_{EU} \times RR_{UD}}{RR_{EU} + RR_{UD} - 1}$ allows for sensitivity analysis on the exposure effect without making any assumption on the unmeasured confounder. Given the lowest bound on the true causal relative risk, can an unmeasured confounder affect the bounding factor in a way to reduce the right-hand side of the above inequality to a value less or equal to 1? This means that an unmeasured confounder leading to such a lower bound of the true causal effect, can explain away the exposure effect on the outcome. Furthermore, if we take $RR_{EU} = RR_{UD}$ then for an unmeasured confounder to explain away the exposure effect on the outcome, it must verify the following equation:

$$RR_{EU} \geq RR_{ED}^{obs} + \sqrt{RR_{ED}^{obs} \times \left(RR_{ED}^{obs} - 1\right)} \tag{11}$$

In general, the two quantities RR_{EU} and RR_{UD} may not be equal but the maximum between the two should satisfy

$$\max(RR_{EU}, RR_{UD}) \geq \frac{RR_{ED}^{obs} + \sqrt{RR_{ED}^{obs} \times \left(RR_{ED}^{obs} - 1\right)}}{RR_{ED}^{\text{true}}} \tag{12}$$

The quantity $RR_{ED}^{obs} + \sqrt{RR_{ED}^{obs} \times \left(RR_{ED}^{obs} - 1\right)}$ is called E-value and we will be investigating how it applies to our results.

5. Results

5.1. Logistic regression

We ran logistic regression each year from 2017 to 2022. In 2017, the value of the property was not available in HMDA data, so it was not possible to derive LTV. Table 2 shows the odds ratios and 2.5% and 97.5% confidence limits of each minority variable. We show the odds ratios for Black, Hispanic and Asian applicants and their confidence intervals (CI) in Table 2 and plot the trends in Fig. 3. Black applicants have the highest odds of denial year over year compared to applicants in other minority groups and this odds ratio has been steady between 2.3 and 2.7 with the upper limit of the confidence interval ranging from 2.73 to 3.23 across the study period. The

Table 2
Logistic regression results by year.

Variable	Year	Odds Ratio	CI 2.5	CI 97.5
Black	2017	2.67	2.21	3.23
Hispanic	2017	1.78	1.37	2.28
Asian	2017	1.24	1.01	1.51
RaceNA	2017	1.92	1.63	2.25
Black	2018	2.56	2.13	3.07
Hispanic	2018	1.82	1.45	2.26
Asian	2018	1.81	1.52	2.16
RaceNA	2018	1.67	1.41	1.97
Black	2019	2.44	2.01	2.95
Hispanic	2019	1.71	1.35	2.16
Asian	2019	2.06	1.73	2.46
RaceNA	2019	1.92	1.62	2.27
Black	2020	2.50	2.08	3.00
Hispanic	2020	1.22	0.94	1.58
Asian	2020	1.78	1.47	2.15
RaceNA	2020	1.68	1.42	1.99
Black	2021	2.30	1.93	2.73
Hispanic	2021	1.46	1.18	1.80
Asian	2021	1.48	1.24	1.77
RaceNA	2021	1.61	1.36	1.89
Black	2022	2.50	2.07	3.03
Hispanic	2022	1.53	1.22	1.90
Asian	2022	1.39	1.15	1.69
RaceNA	2022	1.91	1.61	2.28

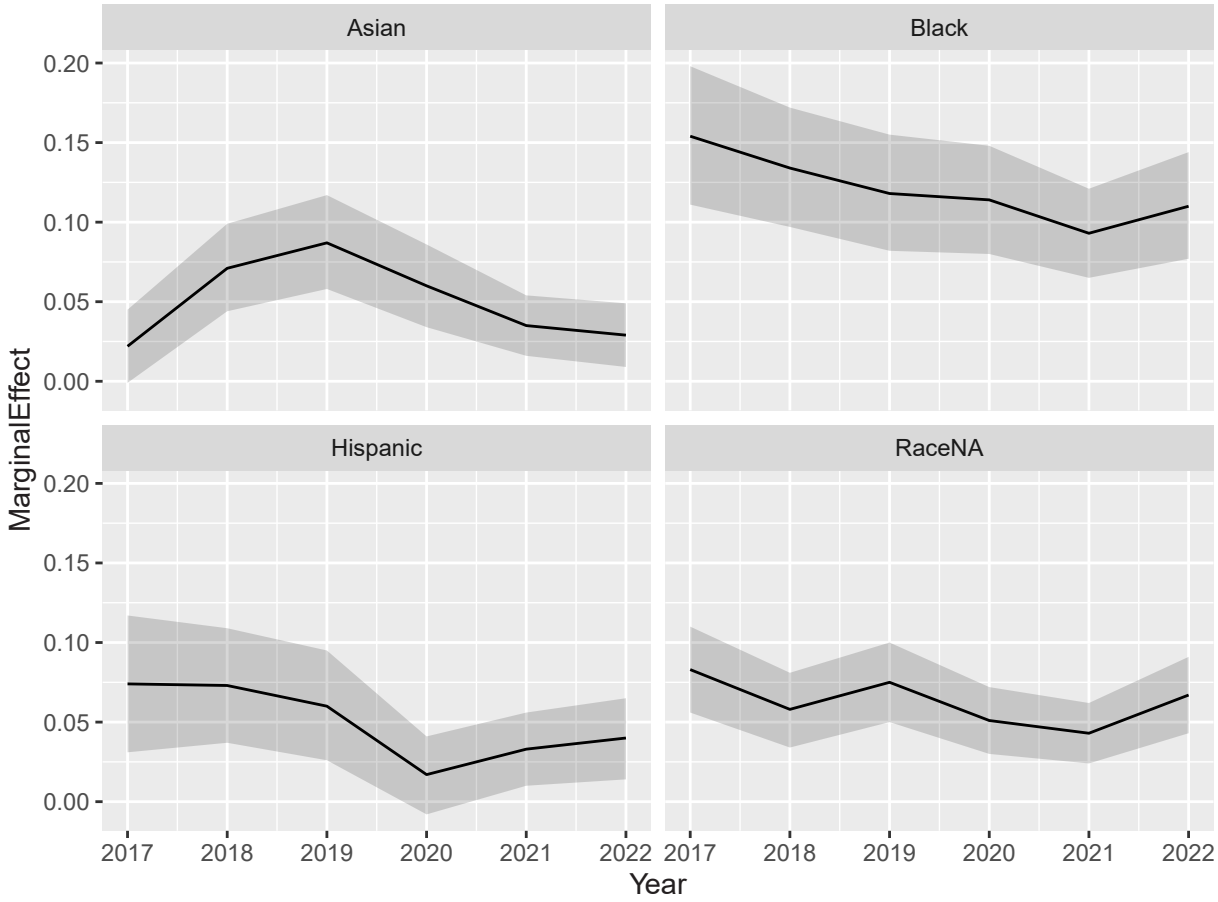


Fig. 3. Trend of odds ratio for minority groups.

odds ratios of the other variables are given in Table 9 in Appendix 3 where we see that among the predictors of denial, the percentage of minority population in tract "tract_minority_population_percent" is the most influential. Each unit increase in the tract minority percentage increases the odds of denial of the loan by 1.97 times in 2018 and up to 2.12 times in 2019.

As noted by the CFPB in their Fall 2015 Supervisory Highlights [30], although odds ratios are commonly used as a metric to assess disparities, "'odds' conveyed by an odds ratio are not the same as probabilities and the magnitudes of odds ratios can be difficult to interpret in some situations. For example, an odds ratio of 3 does not indicate that the denial rate for the minority group is three times the denial rate for the control group." An odds ratio of 3 could result from scenarios such as:

- The minority group is denied in 25% of cases, versus 10% of cases for the control group,
- The minority group is denied in 99% of cases, versus 97% of cases for the control group.

Note that in these scenarios, the prevalence of the event (denial in this case) is not rare, and odds ratios cannot reasonably approximate risk ratios.

Another metric for estimating disparities is the conditional marginal effect (ME), which expresses the increase in probability of denial for a given group, compared to the reference group, while holding all the other variables constant. It is much easier to interpret, and in Table 3, we provide the marginal effect and its 95% confidence interval for each minority group. Here again, we see that the probability of denial for Black applicants has been approximately 15% in 2017 and dropped to approximately 11% in 2022. For Asian applicants, it was 2% in 2017 but not statistically significant (as the 95% confidence interval includes 0), increased to 8.7% in 2019 and decreased each year after that to end at 2.9% in 2022. Hispanic applicants have seen a steady decrease in their marginal effect from 7.4% in 2017 to 1.7% (and not statistically different from 0 in 2020) and then an increase to 4% in 2022. Note that the marginal effect in Table 3 reflects the change in probability of denial if White applicants were in one of the minority groups. For instance, for 2022, the change in marginal effect from 5.8% for White applicants to 11.0% marginal effect for Black applicants, reflects the counterfactual "what would be the change in the denial probability if every White applicant in the data were Black?". The reference group (Minority 0) has different marginal effect depending on the minority group it is compared to; this is because each marginal effect is derived from a regression that involves only one of the minority groups and similarly situated White applicants, hence the regression coefficients

Table 3
Marginal effects by minority group and year.

Year	Minority	ME Minority = 0	LowerCI	UpperCI	ME Minority = 1	LowerCI	UpperCI
2017	Black	0.080	0.066	0.094	0.154	0.111	0.198
2017	Hispanic	0.050	0.029	0.072	0.074	0.031	0.117
2017	Asian	0.019	0.002	0.036	0.022	−0.001	0.045
2017	RaceNA	0.054	0.041	0.066	0.083	0.056	0.110
2018	Black	0.073	0.060	0.086	0.134	0.097	0.172
2018	Hispanic	0.049	0.032	0.067	0.073	0.037	0.109
2018	Asian	0.048	0.035	0.061	0.071	0.044	0.099
2018	RaceNA	0.041	0.028	0.054	0.058	0.034	0.081
2019	Black	0.065	0.052	0.078	0.118	0.082	0.155
2019	Hispanic	0.042	0.024	0.060	0.060	0.026	0.095
2019	Asian	0.054	0.042	0.066	0.087	0.058	0.117
2019	RaceNA	0.048	0.037	0.060	0.075	0.050	0.100
2020	Black	0.059	0.049	0.070	0.114	0.080	0.148
2020	Hispanic	0.015	−0.004	0.033	0.017	−0.008	0.041
2020	Asian	0.040	0.027	0.052	0.060	0.034	0.086
2020	RaceNA	0.035	0.025	0.046	0.051	0.030	0.072
2021	Black	0.051	0.042	0.061	0.093	0.065	0.121
2021	Hispanic	0.025	0.011	0.039	0.033	0.010	0.056
2021	Asian	0.026	0.015	0.037	0.035	0.016	0.054
2021	RaceNA	0.031	0.021	0.041	0.043	0.024	0.062
2022	Black	0.058	0.048	0.069	0.110	0.077	0.144
2022	Hispanic	0.029	0.015	0.044	0.040	0.014	0.065
2022	Asian	0.023	0.010	0.036	0.029	0.009	0.049
2022	RaceNA	0.043	0.032	0.054	0.067	0.043	0.091

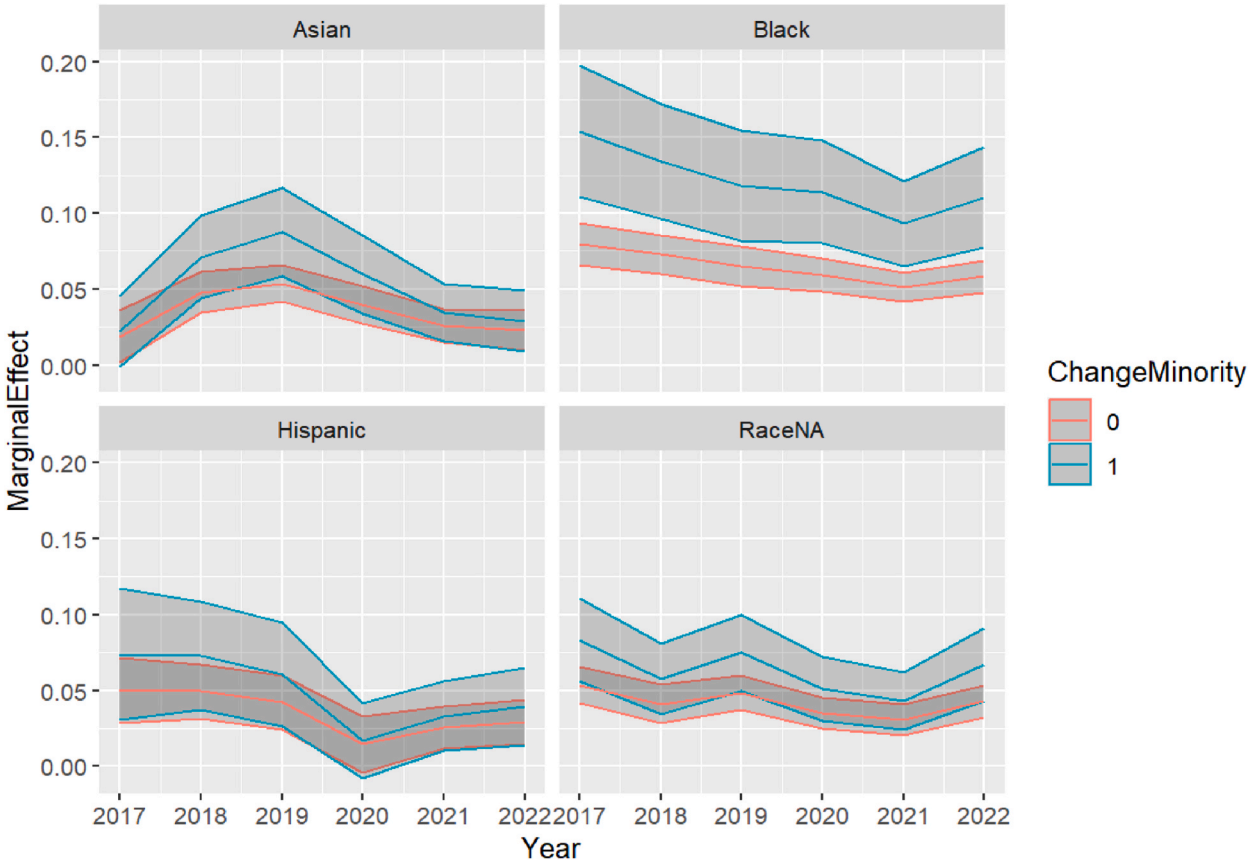


Fig. 4. Trend of marginal effect for minority groups.

differ by minority group. Fig. 4 gives a visual representation of these conditional marginal effects and their 95% confidence interval. Note that for all minority groups, the 95% confidence intervals of the 0 and 1 minority groups largely overlap, however, this is not the case for the Black minority.

One could also report the mean of the change in denial probability, the average marginal effect (AME), for Black (or other minority) applicants. Table 4 gives these quantities for 2022 as well as the 2.5% and 97.5% confidence limits. In that case, the marginal effect for Black applicants is 6.6% (95% confidence interval includes 7%) which agrees with the finding of Giacoletti et al. [26] using national data between 1998 and 2018, despite the difference in time and geography of our data. This reflects a consistent and persistent disparity in underwriting decisions between Black and White applicants. Table 10 in Appendix 4 gives these quantities for all the years.

5.2. Sensitivity analysis

Unmeasured confounders may explain away the significant odds ratio or marginal effect of the minority groups. We know we haven't accounted for FICO, DTI but also others that are difficult to measure such as consumer behavior. These can be masked by the minority groups effects and bias the results of the analysis. To this end, we derive the sensitivity to missing confounders parameters described in Section 4.2.

5.3. Odds ratios

Since the bounding factor applies to the relative risk, we first need to map the odds ratio to a relative risk. Ding and VanderWeele [8] recommend using the odds ratio as an approximation of the relative risk when the event probability is lower than 15%. For each HMDA data from 2017 to 2022, since the denial frequency in Philadelphia metro is below 15% (10.9% in 2017, 10.8% in 2018, 10% in 2019, 8.6% in 2020, 8.3% in 2021 and 8.7% in 2022), we can use the observed odds of loan denial for the minority applicants, controlling for the measured covariates, as an approximation of the observed relative risk. Table 5 shows the odds ratios for minority groups, their 95% confidence bounds and the sensitivity to unmeasured confounder metric as well as its 95% confidence bounds, for each year from 2017 to 2022. The E-value is consistently above 4.0 every year for Black applicants, much higher than for other minority groups, requiring a large bounding factor (equation (10)) to bring the lower bound of the 95% confidence interval of the true odds ratio to 1 or lower, eliminating the disparity in conventional loan denial between Black and White applicants. Take 2017, for Black applicants, the odds ratio of denial is 2.7, the 95% confidence interval is $(2.2, 3.2)$. The E-value for the Black minority group is 4.8 and its 95% confidence interval is (3.8, 5.9). Based on this E-value we can conclude that to explain away the observed relative risk of 2.7, an unmeasured confounder that is associated with the Black minority group and the denial of a loan by a relative risk of 4.8 -fold each, above and beyond the measured confounders, could explain it away, but weaker confounding could not.

The missing E-Value 2.5% confidence limit for Hispanic in 2020 is due to a negative value in the expression of the E-Value $RR_{ED}^{obs} + \sqrt{RR_{ED}^{obs} \times \left(RR_{ED}^{obs} - 1\right)}$ when plugging 0.94, the 2.5% confidence limit of $\mathrm{RR}_{\mathrm{ED}}^{\mathrm{obs}}$. The odds ratio was not statistically significant in that case.

Is an unmeasured confounder such as FICO score strong enough to move the lower limit of the 95% confidence interval of RR_{ED}^{obs} to include 1 for each minority group? We believe this unmeasured confounder associated with denial of loans and with the Black minority group by a relative risk close to 5-fold seems implausible. To illustrate this point, suppose the unmeasured confounder is FICO. Boady et al. [31] published data showing Black and White Americans had a mean FICO score of 677 and 734 respectively, in 2021. Consider the effect of FICO lower than 677 on the denial of conventional home loans and suppose that in 2021, $RR_{UD} = 5$. This translates into:

$$\frac{P[\text{ denial} = 1/FICO < 677, X]}{P[\text{ denial} = 1/FICO >= 677, X]} = 5 \tag{13}$$

Where X refers to other measured covariates we have controlled for. For instance, this could happen if the numerator is 40% and the denominator is 8%, which is not unreasonable since the lowest FICO score required for conventional loan is 620. We can plug this value into equation (14) below to obtain the lowest

bound of RR_{EU} :

$$RR_{EU} \geq \frac{RR_{ED}^{obs} \times RR_{UD} - RR_{ED}^{obs}}{RR_{UD} - RR_{ED}^{obs}} \tag{14}$$

It follows that the smallest value of RR_{EU} needed to neutralize the Black minority group effect on the denial decision obtained in Table 2

Table 4
Average marginal effects by minority group for 2022.

Minority	AME	SE	z	Lower	Upper
Black	0.066	0.007	9.323	0.052	0.080
Asian	0.024	0.007	3.247	0.010	0.038
Hispanic	0.031	0.008	3.769	0.015	0.047
RaceNA	0.046	0.006	7.052	0.033	0.059

Table 5
Odds ratios and sensitivity measures of the Philadelphia Metro disparity analysis over the years.

Variable	Year	Odds Ratio	CI 2.5	CI 97.5	E-Value	E-Value C 2.5	E-Value C 97.5
Black	2017	2.67	2.21	3.23	4.78	3.84	5.90
Hispanic	2017	1.78	1.37	2.28	2.96	2.09	4.00
Asian	2017	1.24	1.01	1.51	1.79	1.12	2.39
RaceNA	2017	1.92	1.63	2.25	3.24	2.64	3.93
Black	2018	2.56	2.13	3.07	4.55	3.67	5.59
Hispanic	2018	1.82	1.45	2.26	3.04	2.26	3.96
Asian	2018	1.81	1.52	2.16	3.03	2.40	3.74
RaceNA	2018	1.67	1.41	1.97	2.72	2.17	3.34
Black	2019	2.44	2.01	2.95	4.31	3.44	5.35
Hispanic	2019	1.71	1.35	2.16	2.82	2.04	3.75
Asian	2019	2.06	1.73	2.46	3.54	2.84	4.35
RaceNA	2019	1.92	1.62	2.27	3.25	2.63	3.97
Black	2020	2.50	2.08	3.00	4.44	3.57	5.46
Hispanic	2020	1.22	0.94	1.58	1.75	NaN	2.54
Asian	2020	1.78	1.47	2.15	2.96	2.31	3.54
RaceNA	2020	1.68	1.42	1.99	2.75	2.20	3.72
Black	2021	2.30	1.93	2.73	4.02	3.28	4.90
Hispanic	2021	1.46	1.18	1.80	2.28	1.63	3.00
Asian	2021	1.48	1.24	1.77	2.33	1.79	2.93
RaceNA	2021	1.61	1.36	1.89	2.59	2.07	3.19
Black	2022	2.50	2.07	3.03	4.45	3.56	5.50
Hispanic	2022	1.53	1.22	1.90	2.42	1.74	3.21
Asian	2022	1.39	1.15	1.69	2.13	1.56	2.77
RaceNA	2022	1.91	1.61	2.28	3.24	2.59	3.99

for 2021, of 2.3, is $RR_{EU} >= 3.40$.

Based on Forbes [20], as of 2021, 5.3% of White Americans have a FICO score below 620

$$P[FICO < 620/\text{Race} = \text{White}] = 0.053 \tag{15}$$

We can use this information to derive the standard deviation of FICO in 2021 for White Americans given their average FICO for that same year is 734. Assuming the distribution of FICO is normal with mean 734 and standard deviation σ, we can conclude that

$$\frac{620 - 734}{\sigma} = -1.62 \tag{16}$$

$\frac{620-734}{\sigma}$ must coincide with the 5.3% quantile of the normal distribution, which is -1.62. We can then derive σ to be equal to 70.38 in 2021 for White Americans. We can derive RR_{EU} for FICO less than 677 as follows:

$$\frac{P[FICO < 677/\text{ Race} = \text{Black }]}{P[FICO < 677/\text{ Race} = \text{White }]} \tag{17}$$

We already have that $P[FICO < 677/\text{ Race} = \text{Black }] = .5$ since the mean FICO is 677 in 2021 for African Americans. All we need to do is normalize the quantile in the denominator based on the mean of 734 and standard deviation of 70.4 for White Americans and we find -0.81. The normal probability corresponding to this quantile is 21%. We can then derive

$$\frac{P[FICO < 677/\text{ Race} = \text{Black }]}{P[FICO < 677/\text{ Race} = \text{White }]} = \frac{.50}{.21} = 2.4 \tag{18}$$

Table 6
Sensitivity measures of the Philadelphia Metro disparity analysis.

Minority	RR ED	E-Value	RR UD	Min RR EU
Black	2.30	4.02	5	3.40
Hispanic	1.46	2.28	5	1.65
Asian	1.49	2.33	5	1.69
RaceNA	1.61	2.59	5	1.89

This value is much smaller than the required minimum of 3.40, which together with the relative risk of FICO of 5 on the denial, would neutralize the Black minority group effect on the denial of 2.3. In fact, in this case, the value of RR_{UD} that leads to $RR_{EU} = 2.4$ is 35, which is extremely large and unlikely to hold for that level of FICO.

Table 6 gives the odds ratios of denial by minority group compared to White applicants as well as the E-value when the RR_{UD} of FICO on the denial is fixed at 5, and the minimum value of RR_{EU} needed to explain away the race effect on the denial of loans for 2021.

Note that we made several assumptions in this illustration, namely that the FICO scores of Black and White mortgage applicants in Philadelphia MSA are similar to those of the Black and White populations in the United States.

Fig. 5 shows the confounder strength frontier for different values of RR_{EU} and RR_{UD}, for Black applicants, in 2021, with the E-value at $(4.0, 4.0)$ shown at the intersection of the red dashed line and the curve.

5.4. Marginal effects

Table 3 reflects the change in the exposure effect if the exposed (minority group) changes from 0 (reference group) to 1 (minority group). Ding and VanderWeele [8] suggest approximating RR_{ED} in the E-value formula by the ratio of marginal effects for exposed and unexposed. From Table 3, the E-value is then derived on the relative risk scale by taking the ratio of ME Minority = 1 to ME Minority = 0. We provide the ratio of these marginal effects and the corresponding E-value by minority and year in Table 7.

For example, the E-value of 3.28 expresses that, to explain away the marginal effect of probability of denial(outcome) for Black applicants(exposure) of 15.4% relative to that of White applicants of 8% in 2017, a confounder with a relative risk of at least 3.28 on the outcome and the exposure is needed.

6. Discussion

The results of this analysis reveal a persistent and statistically robust disparity in conventional mortgage denial rates between Black and White applicants in the Philadelphia MSA from 2017 to 2022, a period that spans pre-pandemic economic conditions, the pandemic-era refinancing boom, and the post-pandemic tightening cycle. The odds of denial for Black applicants remained approximately 2.5 times those of similarly situated White applicants throughout the study period, with the upper bound of the 95% confidence interval at or near three in most years. The conditional marginal effect of approximately 7 percentage points per year is remarkably stable across six years of data and aligns closely with findings from studies covering different time periods and geographies — Munnell and McEneaney [22] found a gap of 8%, Bartlett et al. [25] found 7–10%, and Giacoletti et al. [26] found 7% — despite differences in data vintage, geographic scope, and methodological approach. The consistency of this estimate across decades of research and across different MSAs suggests the disparity reflects structural conditions in the mortgage market rather than transient or geographically specific factors.

The E-value analysis adds a dimension that prior studies of this type have not reported. An E-value consistently above 4.0 for Black applicants throughout the study period means that any unmeasured confounder capable of explaining away the observed disparity would need to be associated with both Black race and loan denial by a relative risk of at least four — above and beyond the covariates already controlled for, which include loan amount, loan-to-income ratio, loan-to-value ratio, income, and neighborhood characteristics. This is a demanding threshold. As the FICO illustration in Section 5.3 demonstrates, even the most plausible and powerful

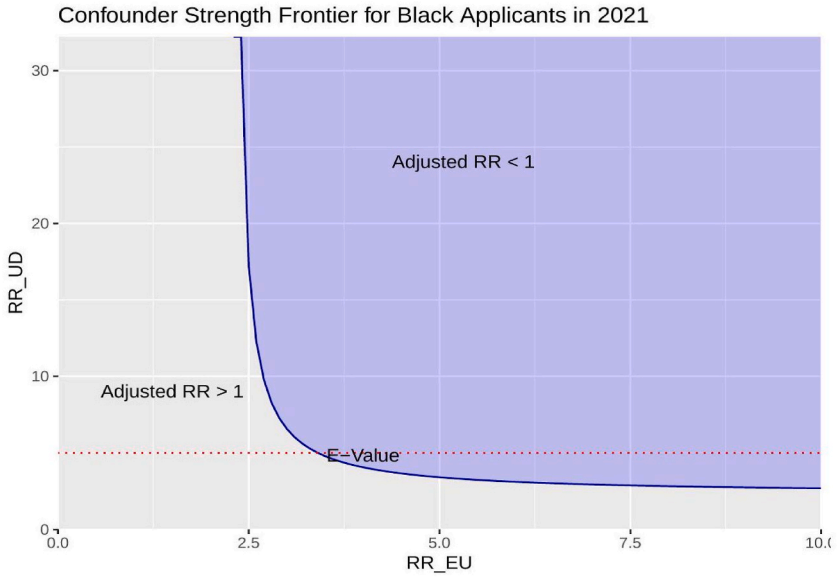


Fig. 5. Confounder Strength Frontier for Black applicants in 2021.

Table 7
Marginal effects and sensitivity measures of the Philadelphia Metro disparity analysis over the years.

Year	Minority	Ratio of Marginal Effects	E-value
2017	Black	1.93	3.28
2017	Hispanic	1.48	2.31
2017	Asian	1.16	1.60
2017	RaceNA	1.55	2.48
2018	Black	1.85	3.09
2018	Hispanic	1.48	2.32
2018	Asian	1.48	2.33
2018	RaceNA	1.40	2.15
2019	Black	1.81	3.03
2019	Hispanic	1.44	2.23
2019	Asian	1.62	2.63
2019	RaceNA	1.55	2.47
2020	Black	1.92	3.25
2020	Hispanic	1.16	1.59
2020	Asian	1.51	2.39
2020	RaceNA	1.45	2.26
2021	Black	1.82	3.03
2021	Hispanic	1.31	1.96
2021	Asian	1.33	2.00
2021	RaceNA	1.41	2.16
2022	Black	1.89	3.19
2022	Hispanic	1.35	2.03
2022	Asian	1.27	1.85
2022	RaceNA	1.56	2.50

omitted confounder — credit score — is unlikely to satisfy it alone, particularly given that the residual denial gap found by Bhutta et al. [24] after controlling for FICO, DTI, and automatic underwriting system results was still 2%. An E-value of 4 applied to Bhutta et al.'s residual gap of 2% produces E-values in the range of 1.7 to 5.4, all of which remain substantial, given that the strongest confounders have already been accounted for.

The contrast between Black applicants and Hispanic and Asian applicants is instructive. Hispanic and Asian denial disparities, while statistically significant in most years, show lower E-values — hovering around 2.0 to 3.0 — indicating greater susceptibility to unmeasured confounding. This does not mean discrimination against Hispanic or Asian applicants is absent; it means the evidence for a causal interpretation is weaker, and that available confounders may account for more of the observed gap. This distinction has practical implications for fair lending examination priorities: the Black-White gap in Philadelphia meets a standard of evidentiary robustness that the Hispanic and Asian gaps do not, at least under the public HMDA data available for this analysis.

The geographic context of Philadelphia reinforces the significance of these findings. The Philadelphia MSA has a well-documented history of residential segregation, redlining, and lending discrimination, and the Federal Reserve Bank of Philadelphia has noted that the African American-White homeownership gap in the city has seen little progress over three decades [3]. The neighborhood-level persistence of HOLC-era lending patterns documented by Namin et al. [15] and Faber's [16] finding that racial segregation shapes perceived creditworthiness in the post-recession mortgage market provide a structural backdrop against which our individual-level results should be interpreted. The disparity we observe is not an artifact of one lender's practices or one year's data — it is a consistent feature of a market with deep institutional roots in racially differentiated lending.

The downward trend in denial disparities for Hispanic and Asian applicants over the study period is encouraging and warrants monitoring. However, the stability of the Black-White gap across the same period, with no meaningful reduction in either the odds ratio or the marginal effect, suggests that market forces and existing regulatory oversight have not been sufficient to close the most persistent dimension of the disparity. From a policy perspective, this finding points toward the need for targeted examination of denial criteria applied to Black applicants, including the threshold levels set for LTV, LTI, and FICO, which lenders calibrate to minimize default risk but which may simultaneously impose disproportionate barriers on minority applicants through facially neutral but discriminatory-in-effect policies — a classic disparate impact mechanism under the Fair Housing Act.

7. Limitations

This paper addresses only one aspect of bias in observational study, namely the unmeasured confounder bias, as it relates to the publicly available HMDA data and the fair lending of underwriting decisions. It does not account for selection bias, measurement errors

and missing data problems. We recognize this is a limitation of our research, but we wanted to focus on one of the most ignored biases, even though it is the central challenge for establishing causality in observational studies and reveal its potential importance to research in the fair lending space. The E-value is a sensitivity metric and less dependent on sample size than the p-value, which expresses a different concept. p-value is often used to assess the evidence of causality in randomized study, however, in observational studies, it is often mis-interpreted and abused, when a presence of association is confused with causation [32,33]. Just like the p-value, the E-value cannot be used to establish a null association. In addition, the scale of the E-value should be considered in relation to potentially known and measured confounders, if these are large, say in the range of 9, then an E-value of 4 could reflect susceptibility of the exposure effect to hidden confounders.

8. Conclusion

This paper shows, through rigorous analysis and several metrics, that disparities in mortgage underwriting are still prevalent especially for African American applicants. Our results are consistent with other authors' findings despite the data being different in time and geography. We quantified the strength of the unmeasured confounder that will explain away the disparity and showed that it will take a very strong unmeasured confounder to accomplish that. It is likely that African Americans are facing disparate impact, possibly through tougher than necessary thresholds on policy variables such as LTV, LTI, FICO or others. Lenders establish these thresholds to minimize the probability of default, but they don't necessarily minimize their fair lending risk at the same time.

CRediT authorship contribution statement

Maia Berkane: Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Diego Alvarez:** Writing – review & editing, Validation, Project administration. **Nico Tanzi:** Writing – review & editing, Visualization, Validation.

Ethics

This study uses publicly available secondary data obtained from the Federal Financial Institutions Examination Council (FFIEC) and the Consumer Financial Protection Bureau (CFPB) through the Home Mortgage Disclosure Act (HMDA) data publication portal (ffiec.cfpb.gov). No primary data was collected and no human subjects were directly recruited, contacted, or involved in this research. The HMDA data are de-identified at the individual level and made publicly available by federal agencies in accordance with statutory disclosure requirements. Accordingly, institutional review board (IRB) or ethics committee approval was not required for this study.

The article is not under consideration for publication elsewhere. The article's publication is approved by all authors. If accepted, the article will not be published elsewhere in the same form, in English or in any other language, including electronically, without the written consent of the copyright holder. Generative AI was not used in the preparation of the manuscript, the research design, the data analysis, or the interpretation of results.

Data availability and materials

The HMDA data used in this study are publicly available from the Federal Financial Institutions Examination Council (FFIEC). Data for 2017 can be accessed from the FFIEC data publication page at: https://ffiec.cfpb.gov/data-publication/three-year-national-loan-level-dataset/2017. Users will need to filter the data to MSA code 37964 (Philadelphia) and apply the additional filters described in Section 4, including loan purpose = 1 (home purchase), before the dataset matches the one used in this study. Data for 2018 to 2022 can be accessed via the FFIEC data browser at: https://ffiec.cfpb.gov/data-browser/data/2018?category=msamds&items=37964&loan_purposes=1, where the year in the link can be changed to any year from 2019 to 2022. The data browser link is pre-filtered to the Philadelphia MSA (code 37964) and home purchase loans, reflecting some of the filters applied in this study. Data filters, manipulations, transformations, and variable constructions are described in detail in Section 4 of the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The study was conducted independently and was motivated by the authors' academic interest in the subject and the objective of contributing to scholarly and public understanding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix 1

The HMDA fields comprise Universal Loan Identifier (ULI), Application Date, Loan Type, Loan Purpose, Occupancy Type, Loan Amount, Action Taken, Action Taken Date, Census Tract, Ethnicity of Applicant or Borrower, Ethnicity of Co-Applicant or Co-Borrower, Race of Applicant or Borrower, Race of Co-Applicant or Co-Borrower, Sex of Applicant or Borrower, Sex of Co-Applicant or Co-Borrower, Age of Applicant or Borrower, Age of Co-Applicant or Co-Borrower, Income, Lien Status, Credit Score of Applicant or Borrower, Credit Score of Co-Applicant or Co-Borrower, Origination Charges. Discount Points, Lender Credits, Interest Rate, Debt-to-Income Ratio, Combined Loan-To-Value Ratio, Loan Term, Property Value, Manufactured Home Secured Property Type, Total Units, Automated Underwriting System Result, Reverse Mortgage, Open-End Line of Credit, Business or Commercial Purpose.

To standardize the definition of race and ethnicity, MSA, loan product and dwelling category, CFPB provides its derived variables as well as census data related to the MSA or metro division (MD) of each application. CFPB derived HMDA fields include derived race, derived ethnicity, derived sex, derived loan product type, derived dwelling category, conforming loan limit, derived msa md, applicant age above 62, co-applicant age above 62, tract population, tract minority population percent, FFIEC msa_md median family income, tract owner occupied units, tract one-to-four family homes, tract median age of housing units.

Appendix 2

Table 8
Mean Differences in Key Variables Between Minority and White Applicants (t-tests with 95% Confidence Intervals)

Variable	Minority	Estimate	Minority Mean	White Mean	Statistic	p-value	df	95% CI (Lower)	95% CI (Upper)
ln_LoanAmount	Black	−0.6540	11.948	12.602	−28.685	1E-157	3E+03	−0.69869	−0.60928
ln_LoanAmount	Hispanic	−0.6557	11.947	12.602	−24.189	2E-110	2E+03	−0.70886	−0.60252
ln_LoanAmount	Asian	−0.1045	12.498	12.602	−6.770	2E-11	3E+03	−0.13476	−0.07423
ln_LoanAmount	RaceNA	−0.0916	12.511	12.602	−4.977	7E-07	4E+03	−0.12766	−0.05550
ln_Income	Black	−0.5709	11.109	11.680	−38.666	4E-272	3E+03	−0.59982	−0.54193
ln_Income	Hispanic	−0.6945	10.986	11.680	−38.119	8E-236	2E+03	−0.73020	−0.65874
ln_Income	Asian	−0.3286	11.352	11.680	−19.690	3E-81	3E+03	−0.36128	−0.29584
ln_Income	RaceNA	−0.0511	11.629	11.680	−3.331	9E-04	4E+03	−0.08119	−0.02103
tract_minority_population_percent	Black	0.4083	0.728	0.320	61.067	0E+00	3E+03	0.39516	0.42138
tract_minority_population_percent	Hispanic	0.3326	0.652	0.320	41.697	4E-259	2E+03	0.31692	0.34821
tract_minority_population_percent	Asian	0.1421	0.462	0.320	22.378	6E-102	3E+03	0.12965	0.15456
tract_minority_population_percent	RaceNA	0.1139	0.434	0.320	19.133	6E-78	4E+03	0.10223	0.12557
tract_to_msa_income_percentage	Black	−0.6615	0.939	1.601	−53.516	0E+00	4E+03	−0.68570	−0.63724
tract_to_msa_income_percentage	Hispanic	−0.6729	0.928	1.601	−42.310	3E-283	2E+03	−0.70408	−0.64170
tract_to_msa_income_percentage	Asian	−0.3079	1.293	1.601	−20.039	3E-84	3E+03	−0.33806	−0.27780
tract_to_msa_income_percentage	RaceNA	−0.1341	1.467	1.601	−8.835	1E-18	4E+03	−0.16385	−0.10434
LTI	Black	0.1679	3.041	2.873	4.262	2E-05	3E+03	0.09063	0.24510
LTI	Hispanic	0.4647	3.338	2.873	9.773	6E-22	2E+03	0.37145	0.55798
LTI	Asian	0.7043	3.577	2.873	17.493	6E-65	3E+03	0.62534	0.78323
LTI	RaceNA	−0.0042	2.869	2.873	−0.136	9E-01	4E+03	−0.06515	0.05668
LTV	Black	0.0175	0.820	0.802	2.815	5E-03	3E+03	0.00532	0.02976
LTV	Hispanic	0.0096	0.812	0.802	1.325	2E-01	2E+03	−0.00463	0.02392
LTV	Asian	−0.0227	0.780	0.802	−5.359	9E-08	4E+03	−0.03100	−0.01439
LTV	RaceNA	0.0017	0.804	0.802	0.370	7E-01	4E+03	−0.00732	0.01073
tract_owner_occupied_units	Black	−107.7877	1008.602	1116.389	−9.611	1E-21	3E+03	−129.77607	−85.79940
tract_owner_occupied_units	Hispanic	−35.2627	1081.127	1116.389	−2.535	1E-02	2E+03	−62.54232	−7.98309
tract_owner_occupied_units	Asian	−39.0510	1077.338	1116.389	−3.151	2E-03	3E+03	−63.34892	−14.75306
tract_owner_occupied_units	RaceNA	−73.2560	1043.133	1116.389	−6.963	4E-12	5E+03	−93.88089	−52.63110

Appendix 3

Table 9
Logistic Regression Results by Year

Variable	Year	Odds Ratio	CI 2.5	CI 97.5
ln_Income	2017	1.203925734	0.972181913	1.488029607
ln_LoanAmount	2017	0.566841404	0.485406424	0.664063713
LTI	2017	1.363201422	1.240506221	1.496663342
CoappN	2017	1.046465388	0.910614198	1.203987044
tract_minority_population_percent	2017	1.765384161	1.357501999	2.297352389

(*continued on next page*)

Table 9 (*continued*)

Variable	Year	Odds Ratio	CI 2.5	CI 97.5
tract_to_msa_income_percentage	2017	1.159842053	1.019061279	1.318508121
ln_Income	2018	0.56625281	0.460240537	0.695687864
ln_LoanAmount	2018	0.947148335	0.797694907	1.126482961
LTI	2018	1.086075477	0.997815451	1.181195826
LTV	2018	0.270658055	0.187297723	0.391984163
CoappN	2018	0.931083459	0.814065966	1.066056207
tract_minority_population_percent	2018	1.973332533	1.530063459	2.547254141
tract_to_msa_income_percentage	2018	1.131755711	0.985366168	1.298127598
ln_Income	2019	1.312794611	1.021981258	1.683113412
ln_LoanAmount	2019	0.418102497	0.330575779	0.529409947
LTI	2019	1.449001675	1.317896542	1.592189276
LTV	2019	0.414752686	0.281894579	0.611305364
CoappN	2019	0.99082019	0.862105489	1.140285642
tract_minority_population_percent	2019	2.124148338	1.638577445	2.756925847
tract_to_msa_income_percentage	2019	1.291436802	1.121969334	1.484612889
ln_Income	2020	0.676323829	0.541207015	0.844406329
ln_LoanAmount	2020	0.753612171	0.62772468	0.905977022
LTI	2020	1.178921848	1.090678607	1.273639591
LTV	2020	0.431537103	0.280457914	0.66768908
CoappN	2020	0.782988467	0.67957543	0.903166873
tract_minority_population_percent	2020	2.49122609	1.924054016	3.229484654
tract_to_msa_income_percentage	2020	1.269059864	1.102249118	1.458903617
ln_Income	2021	0.851608705	0.688803026	1.052621922
ln_LoanAmount	2021	0.739025282	0.623050059	0.876774759
LTI	2021	1.350164554	1.266792321	1.438867703
LTV	2021	0.319506291	0.211047864	0.48562234
CoappN	2021	0.888090299	0.775641234	1.018277591
tract_minority_population_percent	2021	1.857684198	1.459695571	2.366196734
tract_to_msa_income_percentage	2021	0.998342786	0.868844952	1.14509637
ln_Income	2022	0.800733968	0.640235298	1.000598961
ln_LoanAmount	2022	0.868404986	0.728377629	1.035785902
LTI	2022	1.370491588	1.273018639	1.475014364
LTV	2022	0.171287752	0.112706762	0.261177305
CoappN	2022	1.109003835	0.958797823	1.285158813
tract_minority_population_percent	2022	2.089310379	1.551930038	2.816504556
tract_to_msa_income_percentage	2022	1.089848042	0.939947605	1.261729525

Appendix 4

Table 10
Average Marginal Effects by Minority Group and Year

Year	Minority	AME	se	z	lower	upper
2017	Black	0.086	0.009	10.041	0.069	0.103
2017	Asian	0.024	0.009	2.580	0.006	0.042
2017	Hispanic	0.052	0.012	4.494	0.029	0.074
2017	RaceNA	0.060	0.007	8.034	0.045	0.074
2018	Black	0.079	0.008	9.704	0.063	0.095
2018	Asian	0.051	0.008	6.568	0.036	0.067
2018	Hispanic	0.051	0.010	5.121	0.031	0.070
2018	RaceNA	0.044	0.007	5.880	0.029	0.058
2019	Black	0.072	0.008	9.042	0.056	0.087
2019	Asian	0.059	0.007	8.051	0.045	0.073
2019	Hispanic	0.043	0.010	4.476	0.024	0.062
2019	RaceNA	0.053	0.007	7.586	0.039	0.066
2020	Black	0.065	0.007	9.263	0.051	0.078
2020	Asian	0.042	0.007	5.867	0.028	0.056
2020	Hispanic	0.014	0.010	1.435	−0.005	0.033

(*continued on next page*)

Table 10 (*continued*)

Year	Minority	AME	se	z	lower	upper
2020	RaceNA	0.038	0.006	6.035	0.026	0.051
2021	Black	0.057	0.006	9.055	0.044	0.069
2021	Asian	0.027	0.006	4.280	0.015	0.040
2021	Hispanic	0.026	0.008	3.417	0.011	0.041
2021	RaceNA	0.033	0.006	5.519	0.021	0.044
2022	Black	0.066	0.007	9.323	0.052	0.080
2022	Asian	0.023	0.007	3.248	0.009	0.038
2022	Hispanic	0.031	0.008	3.769	0.015	0.047
2022	RaceNA	0.046	0.007	7.052	0.033	0.059

References

[1] U.S. Department of Justice. Equal Credit Opportunity Act. Available from: https://www.justice.gov/crt/equal-credit-opportunity-act-3.
[2] L. Quillian, J.J. Lee, B. Honoré, Racial discrimination in the US housing and mortgage lending markets: a quantitative review of trends, 1976-2016, Race and Social Problems 12 (2020) 13–28.
[3] J. Whiton, T. Singleton, L. Ding, What's Behind the Racial Homeownership Gap in Philadelphia? Federal Reserve Bank of Philadelphia, 2021.
[4] U.S. Department of Justice, Office of Public Affairs, Justice department secures over $3 million redlining settlement involving ESSA Bank & Trust in Philadelphia, Available from: https://www.justice.gov/archives/opa/pr/justice-department-secures-over-3-million-redlining-settlement-involving-essa-bank-trust, 2023.
[5] Consumer Financial Protection Bureau, CFPB and DOJ Order Trident Mortgage Company to Pay More than $22 Million for Deliberate Discrimination Against Minority Families, 2022. Available from: https://www.consumerfinance.gov/about-us/newsroom/cfpb-doj-order-trident-mortgage-company-to-pay-more-than-22-million-for-deliberate-discrimination-against-minority-families/.
[6] Berkane, M. (submitted for publication) "Robust Benchmarking for Discrimination Detection in Heterogeneous Markets". Manuscript under review.
[7] Federal Financial Institutions Examination Council, HMDA data, data available from: https://ffiec.cfpb.gov/data-publication/three-year-national-loan-level-dataset/2017.2018–2022dataavailablefrom, 2017 https://ffiec.cfpb.gov/data-browser/data/2018?category=msamds&items=37964&loan_purposes=1.
[8] P. Ding, T.J. VanderWeele, Sensitivity analysis without assumptions, Epidemiology 27 (3) (2016) 368377.
[9] M.M. Brooks, Persistent disparities in affordable rental housing among America's ethno-racial groups, Soc. Sci. Res. 113 (2023) 102828.
[10] M. Kuebler, J.S. Rugh, New evidence on racial and ethnic disparities in homeownership in the United States from 2001 to 2010, Soc. Sci. Res. 42 (5) (2013) 1357–1374.
[11] R. Sanchez-Moyano, Geography and Hispanic homeownership: a review of the literature, J. Hous. Built Environ. 36 (2021) 215–240.
[12] B. Wei, F. Zhao, Racial disparities in mortgage lending: new evidence based on processing time, SSRN Electron. J. 11 (3) (2022) 775–813.
[13] J. Loya, Ethno-racial and relative income disparities in access to mortgage credit, Res. Soc. Stratif. Mobil. 85 (2023) 100785.
[14] J. Loya, Ethno-racial stratification in the mortgage market: the role of co-applicants, Soc. Sci. Res. 106 (2022) 102725.
[15] S. Namin, Y. Zhou, W. Xu, E. McGinley, C. Jankowski, P. Laud, K. Beyer, Persistence of mortgage lending bias in the United States: 80 years after the Home Owners' Loan Corporation security maps, Journal of Race, Ethnicity and the City 3 (1) (2022) 70–94.
[16] J.W. Faber, Segregation and the geography of creditworthiness: racial inequality in a recovered mortgage market, Housing Policy Debate 28 (2) (2018) 215–247.
[17] D.J. Hammel, I. Nilsson, Mortgage foreclosures, race, and postrecession lending, Prof. Geogr. 71 (3) (2019) 536–550.
[18] T. Haupert, New technology, old patterns: fintech lending, metropolitan segregation, and subprime credit, Race and Social Problems (2022) 1–15.
[19] T. Haupert, Racial patterns in mortgage lending outcomes during and after the subprime boom, Housing Policy Debate 29 (6) (2019) 947–976.
[20] Forbes Advisor, From Inherent Racial Bias to Incorrect Data: the Problems with Current Credit Scoring Models, 2021. Available from: https://www.forbes.com/advisor/credit-cards/from-inherent-racial-bias-to-incorrect-data-the-problems-with-current-credit-scoring-models/.
[21] Urban Institute, Credit health during the COVID-19 pandemic, Available from: https://apps.urban.org/features/credit-health-during-pandemic/, 2023.
[22] A.H. Munnell, G.M. T, L.E. B, J. McEneaney, Mortgage lending in Boston: interpreting HMDA data, Am. Econ. Rev. 86 (1) (1996) 25–53.
[23] S.L. Ross, J. Yinger, The Color of Credit, The MIT Press, 2002.
[24] N. Bhutta, A. Hizmo, D. Ringo, How much does racial bias affect mortgage lending? Evidence from human and algorithmic credit decisions, SSRN Electron. J. (2025).
[25] R. Bartlett, A. Morse, R. Stanton, N. Wallace, Consumer-lending discrimination in the FinTech era, J. Financ. Econ. 143 (1) (2022) 30–56.
[26] M. Giacoletti, R. Heimer, E.G. Yu, Using high-frequency evaluations to estimate discrimination: evidence from mortgage loan officers, SSRN Electron. J. (2021).
[27] P.R. Rosenbaum, D.B. Rubin, Assessing sensitivity to an unobserved binary covariate in an observational Study with binary outcome, J. Roy. Stat. Soc. B Stat. Methodol. 45 (2) (1983) 212–218.
[28] D.Y. Lin, B.M. Psaty, R.A. Kronmal, Assessing the sensitivity of regression results to unmeasured confounders in observational studies, Biometrics 54 (3) (1998) 948.
[29] Federal Financial Institutions Examination Council. Loan application register data fields documentation. Available from: https://ffiec.cfpb.gov/documentation/publications/loan-level-datasets/lar-data-fields.
[30] Consumer Financial Protection Bureau, Supervisory highlights, fall, Available from: https://files.consumerfinance.gov/f/201510_cfpb_supervisory-highlights.pdf, 2015.
[31] K. Boady, M. McComas, A. Ouazad, An Analysis of Financial Institutions in black-majority Communities: Black Borrowers and Depositors Face Considerable Challenges in Accessing Banking Services, Brookings Institute, 2021.
[32] S. Greenland, S.J. Senn, K.J. Rothman, J.B. Carlin, C. Poole, S.N. Goodman, D.G. Altman, Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations, Eur. J. Epidemiol. 31 (4) (2016) 337–350.
[33] S. Greenland, Null misinterpretation in statistical testing and its impact on health risk assessment, Prev. Med. 53 (4–5) (2011) 225–228.