
Photo: fizkes/iStock

Banking Trends

Discrimination in Mortgage Markets

Automated underwriting may reduce but likely does not end discrimination against racial minorities.

Edison Yu
Economic Advisor
and Economist
FEDERAL RESERVE BANK
OF PHILADELPHIA

The views expressed in this article are not necessarily those of the Federal Reserve.

Racial discrimination has haunted mortgage markets for decades, prompting legislation and public policy debates that continue to shape how all of us get our mortgages. A new technology may help reduce this discrimination.

Previous research into mortgage markets has shown evidence of a century's worth of racial discrimination, including redlining, unequal mortgage access, and differences in mortgage costs. Federal, state, and local governments have responded to this discrimination by enacting laws such as the federal Fair Housing Act of 1968.[1]

These laws are not relics of a vanished era. Many banks were recently fined a large amount of money due to antidiscrimination lawsuits. For example, in 2019, Wells Fargo Bank wrote the City of Philadelphia a check for $10 million to settle a lawsuit alleging that the bank engaged in discriminatory lending practices.[2]

Some analysts argue that algorithmic or automated underwriting (AU), which has become increasingly popular in mortgage markets, renders lenders less likely to discriminate because it does not use race as an input and presumably bases loan decisions only on the applicant's financial data, limiting the discretionary judgement of human decisionmakers. If that's the case, then AU may help antidiscrimination efforts. This is why we need to study AU's impact on the mortgage lending business.

Researchers disagree as to how best to measure racial discrimination in the mortgage markets, so, before studying AU's impact on antidiscrimination efforts, I survey the history of the statistical methods used to identify racial discrimination. Although it may seem straightforward, identifying racial discrimination is challenging since researchers cannot observe all the information used by loan officers or borrowers. I then describe how, in a previously published working paper, my coauthors and I addressed this challenge by using high-frequency data. I conclude by exploring preliminary evidence on the effects of AU on racial discrimination in the mortgage markets.

The Mortgage Application Process

In the U.S, a mortgage application typically starts with a borrower contacting a potential lender to inquire about mortgage products. The borrower usually has an initial conversation with a loan officer, who is the front person within the lender organization responsible for communicating with the borrower. The loan officer can gauge the borrower's initial eligibility by using the borrower's basic financial information, such as the loan amount needed and credit scores, and guide the borrower to select a mortgage product. Once a mortgage product is selected, the borrower can submit a formal mortgage application.

Increasingly in recent years, borrowers have contacted potential lenders over the Internet.[3] Borrowers search for potential mortgage products on mortgage-shopping platforms such as Zillow or with fintech lenders such as Quicken Loans.[4] After the initial search, a borrower is contacted by the lender, usually by the lender's loan officer. Even at fintech lenders, human loan officers are involved in the application process.

The loan officer's involvement does not end with the initial contact. After a borrower submits a formal application, the loan officer ensures that the borrower has submitted all the necessary documentation, such as verification of income and employment, credit reports, and property appraisal reports. The officer also ensures the accuracy of the information in the application. Once the application is complete, the loan officer sends it to a loan underwriter, who makes the final credit decision based on the application and supporting documents. Although loan officers do not make final credit decisions, they influence the potential credit decision by nudging the borrower to provide timely and accurate documentation. Thus, a loan officer's racial bias can still affect the outcome of a mortgage application. For example, a loan officer might not inform a minority applicant of an incomplete application in a timely manner, leading to rejection of the application by the underwriter.

AU may help antidiscrimination efforts. This is why we need to study AU's impact on the mortgage lending business.

Many lenders increasingly rely on AU, which was used in about 56 percent of mortgage applications in 2019. An AU system processes an applicant's financial information and recommends whether to approve the loan. This recommendation is generated by a computer, not a human. These underwriting systems do not use race as an input and presumably base loan decisions only on the applicant's financial data. Antidiscrimination regulations allow lenders, when making loan decisions, to use variables directly related to credit history and risks, but they prohibit lenders from discriminating on the basis of race, religion, color, national origin, sex, familial status, or disability status.[5]

However, there are a few reasons why the final credit decision might still be biased. Algorithms may produce decisions that unintentionally correlate with impermissible variables (such as race and gender).[6] Since the algorithms can be quite complicated, decision makers may not understand that the algorithms are making biased underwriting choices.[7] In addition, a loan officer's bias may affect the timeliness and completeness of the information input into the AU system. Finally, lenders do not completely rely on the AU-generated recommendation; a human underwriter still makes the final loan decision. AU decisions are only recommendations. So, the final loan decision may still be biased.

Identification of Racial Discrimination Using Observational Data

Social scientists have documented racial disparities in a wide range of areas, including in labor markets, credit markets, and the legal system. Because of the limitations of empirical tests, social scientists disagree as to whether or not these disparities are the result of discrimination by economic decision makers.

The two main types of tests used to identify racial discrimination are benchmarking tests and outcome tests. Benchmarking tests (also known as audit tests) use observational data in a straightforward way to test for discrimination. In the context of mortgage lending, if, after adjusting for credit risks, a minority group receives a lower average approval rate or higher interest rate for the same mortgage product, then the test has identified discrimination. Benchmarking tests are useful because they can be executed in real time.[8] However, benchmarking tests are vulnerable to omitted-variable bias. The omitted variables are the differences in group characteristics that the researcher does not observe but that can cause differences in evaluations. For example, minority applicants might have riskier financial profiles, and the approval rate gap seen in the data could reflect differences in credit risk rather than discrimination.

One common approach for solving this problem is to include additional variables as control variables in the analysis. For example, one can look at the mortgage approval gap when comparing minority and majority applicants while controlling for financial variables such as income and credit risk. However, no researcher has data on all the variables observed by decision makers or borrowers. For example, researchers may not have the "soft" information that loan officers have from their interactions with borrowers. In antidiscrimination lawsuits where minority applicants charge that they received a higher interest rate or larger fees, an oft-used counterargument is that minority applicants do not shop around as much as majority applicants. That kind of data is not generally available to researchers, regulators, or even lenders. So, although it helps to include more control variables, they cannot eliminate the omitted-variable bias problem.

An alternative is to use an outcome test. Instead of comparing differences in how groups are evaluated, such as approval rates, outcome tests compare the subsequent performance of successful applicants through, for example, default rates.[9] Suppose there is a cutoff of application quality, below which the application will be rejected. The applicant who just barely meets the cutoff is the marginal applicant. If there is discrimination, the minority group will face a higher threshold for inclusion, and the marginal minority applicant will thus have better ex post outcomes (for example, a lower default rate) than the marginal majority applicant.

Though intuitively appealing, outcome tests are notoriously difficult to implement. We rarely have data that enable us to identify the marginal applicant, so researchers usually use average quantities (such as the average default rate) instead. But the average difference in ex post outcomes can be a poor approximation of the difference in marginal outcomes. Since we observe only average performance, an observation that the default rate among African Americans is the same as among whites, for example, should not be interpreted as evidence of nondiscrimination.[10] Studies using this approach are less likely to find evidence of discrimination.[11] Thus, implementing an outcome test is very challenging, and many papers in the literature have pointed out its limitations.[12] For that reason, I focus on benchmarking tests.

Evidence of Racial Discrimination in the Mortgage Markets

There is a long history of attempts to identify racial discrimination in the mortgage markets using observational data. Many of these papers try to solve the omitted-variable problem by including more control variables in the statistical analysis.

This literature can be traced back at least as far as the 1996 work of the former director of research at the Federal Reserve Bank of Boston, Alicia Munnell, and her coauthors. In that paper, the authors use the 1990 Home Mortgage Disclosure Act (HMDA) data, but they also use information from a survey that collects additional information from lending institutions operating in the Boston metropolitan statistical area.[13] This additional information includes financial, employment, and property characteristics relevant to a lending decision but missing from the HMDA data. The survey sample covers all applications for conventional mortgage loans made by African American and Hispanic American applicants and a random sample of 3,300 applications made by white applicants. When using the HMDA data alone, the paper finds that the rejection rate of minority applicants is 18 percentage points higher than that of white applicants. When the researchers controlled for the additional information from the survey, the disparity between the rejection rates of minority and white applicants declined to just over 8 percentage points. These results show the importance of controlling for relevant variables absent from the HMDA data set. Still, the rejection rate gap remains large even after adding the controls. Many early papers in the literature show similar results.[14]

In a recent paper, University of California, Berkeley, law professor Robert Bartlett and his coauthors examine whether African Americans and Hispanic Americans pay higher mortgage interest rates than white Americans, and whether this pricing differential remains when the origination is automated. To address the omitted-variable problem, they merged 2009–2015 HMDA data with other data sets that include information about interest rates, the names of lenders, and loan performance. In addition, by using a sample of mortgages insured by a government-sponsored enterprise (GSE), they filter out the default and prepayment risks borne by the lenders. Thus, any disparity in the interest rates paid by minority and majority borrowers should reflect racial discrimination, not credit risk. They find that Hispanic American and African American borrowers collectively pay an additional 7.9 and 3.6 basis points in interest rates for, respectively, purchase mortgages and refinance mortgages. They also find a 40 percent lower level of price discrimination if the lender is a fintech firm, which suggests that AU reduces but does not eliminate discrimination.

However, in their 2021 article, Federal Reserve economists Neil Bhutta and Aurel Hizmo account for more pricing variables, such as discount points and fees, and find no evidence that minorities pay more in mortgages. They supplement the 2014–2015 HMDA data with administrative data from the Federal Housing Administration (FHA) on all FHA-insured mortgages, and with information on points and fees from Optimal Blue, a leading provider of secondary marketing solutions and data services in the mortgage industry. They find a statistically significant gap in interest rates paid by race, but the gap is offset by differences in discount points. They argue that the differences in interest rates across racial groups found in the earlier papers are a result of African American borrowers choosing mortgage products with higher interest rates but lower points, potentially because minority borrowers may find it more difficult to put funds up front. This finding is restricted to FHA mortgages; whether the results generalize to other samples, such as the more GSE-dominated sample used by Robert Bartlett and his coauthors, remains unclear. Notably, lenders who provide FHA mortgage products tend to serve lower-income and minority communities and thus may be less biased.

In another recent paper, Penn State professor of real estate Brent Ambrose and his coauthors find that pricing disparities in mortgage contracts are influenced by whether the borrower and broker are of the same race. They used a novel data set that covers all mortgages approved and funded by New Century Financial Corporation, a now-defunct real estate investment trust, between 2003 and 2007. These loans are representative of the overall subprime market before the Great Recession. This data set comprises more than 300,000 mortgages originated by 124,736 individual brokers, and it contains a rich set of control variables. In addition, the data set includes the names of the brokers. When the authors used a surname-geocoding algorithm to infer each broker's race, they found that minorities pay more in fees than similarly qualified whites, but the premium paid by minorities depends on whether the broker shares their race. For example, African American borrowers who obtain a loan through a white mortgage broker pay 14 percent more than white borrows who work with a white broker, but this premium is lowered to 6 percent when the broker is African American.

For a 2021 working paper, Neil Bhutta and his coauthors studied the impact of AU on racial discrimination using a newly available data set from HMDA. This 2018–2019 HMDA data set provides a longer list of variables—including credit scores, debt-to-income ratios, and AU recommendations—than did earlier HMDA data sets. When they focused on the sample of loans that utilized the AU system, they found that, by controlling for AU recommendations and using these new loan-level variables, the racial gap in mortgage denial rates fell to about 1–2 percentage points. They argue that the remaining gap might be explained by unobserved characteristics of the borrowers, which suggests a more limited role for racial discrimination in mortgages that use AU.

All of these researchers find that the racial gap in mortgage approval rates and costs is very large in the data, but some of this gap can be explained by factors such as credit risk. The question is whether the remaining racial gap is caused by racial bias or by insufficient control of omitted variables. Many of these papers attempt to reduce the problem of omitted variables by adding control variables to the analysis. However, it is difficult to know whether the additional variables eliminate the bias. In addition, these and other papers use samples across different data sets or cover different time periods. This makes it difficult to compare results.

Testing for Discrimination Using High-Frequency Data

In a recent research paper, my coauthors and I took a different approach to address the challenges of identifying racial discrimination.[15] With some assumptions, this approach avoids the problem of omitted variables by employing high-frequency data. We used time variation in loan officers' loan approval decisions to draw inferences about likely discriminatory behavior. We also used the entire HMDA data set from 1994 through 2019, which covers most mortgage applications in the U.S. during those 25 years, making our sample more comprehensive than samples used in earlier work. After discussing this new approach, I will show how we used this new approach to ascertain the impact of AU on discrimination.

First, we find that the volume of mortgage originations increases over the course of a calendar month (Figure 1).[16] The number of loans originated on the last day of a month is almost twice as high as on the first day of the month. There is no similar pattern in application volume. This bunching pattern in originations is likely caused by loan officers' incentive to meet their month-end quota. Loan officers tend to receive a commission that equals a percentage of the total dollar amount they originate during the month. They can also receive a bonus for meeting their monthly origination target. Loan officers who fail to meet volume targets can be disciplined and risk getting fired. Our key insight is that this pressure to meet month-end quotas makes it costlier for loan officers to discriminate at the end of a month.

At the same time, we observe that the mortgage approval gap between white and African American applicants shrinks over the course of the month (Figure 2). In the first seven days of a calendar month, the approval rate gap is close to 20 percent. The gap shrinks in the last days of the month and reaches the lowest point of around 10 percent on the last day of the month. When we control for many observable variables, the approval rate gap shrinks to almost zero on the last day of the month (Figure 3).

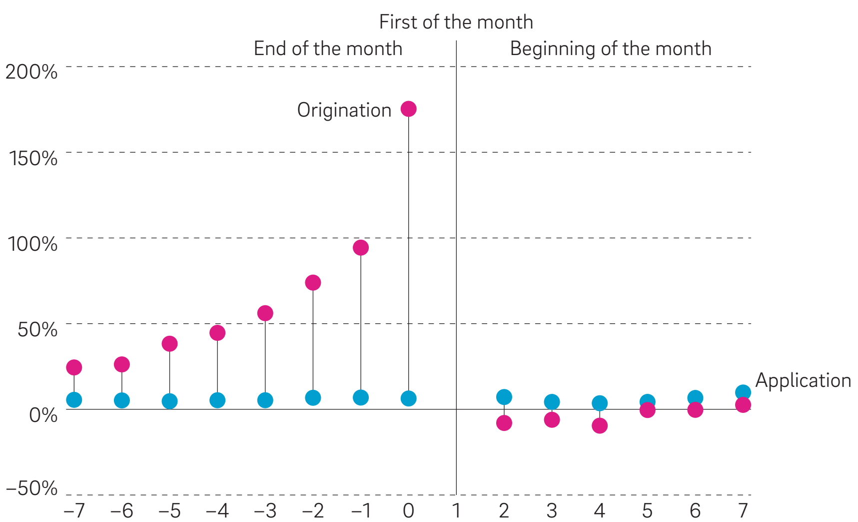
The higher-frequency daily data help us address the omitted-variable bias. In our paper, we discuss a number of potential omitted variable issues. The reduction in the approval gap within a month, as seen in Figure 3, might be attributed to an unobserved within-month movement of application quality rather than changes in discrimination. For example, the gap would be explained without reference to discrimination if the quality of African American applications is higher toward

FIGURE 1

Originations Surge as the Month Ends

Loan officers approve more applications toward the end of the month, most likely to meet their monthly quotas.

Daily loan applications and originations, as a percentage of loan applications and originations on the first day of the month, 1994–2019



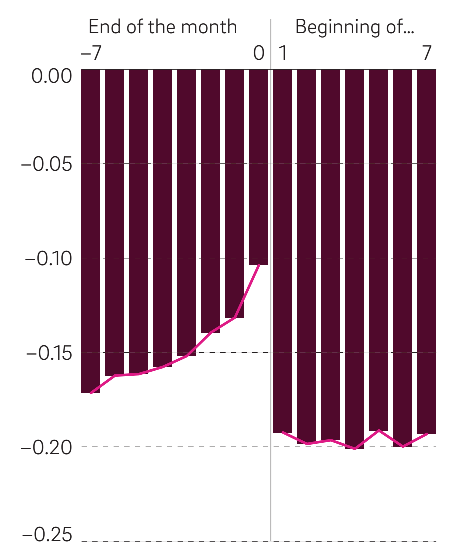
Source: Home Mortgage Disclosure Act (HMDA) data set, Board of Governors of the Federal Reserve System.

Note: The number 0 on the horizonal axis indicates the last day of a calendar month, the positive numbers indicate the first seven days of the month, and the negative numbers indicate the last days of the month.

FIGURE 2

The Approval Rate Gap Shrinks Toward the End of the Month

Approval rate for African American applicants minus rate for white applicants in the seven days preceding and succeeding the first of the month, 1994–2019



Source: Home Mortgage Disclosure Act (HMDA) data set, Board of Governors of the Federal Reserve System.

Note: This figure does not consider observable factors that might affect the approval decision.

the end of a month.[17] But we do not see a within-month bunching pattern in application composition and observed applicant quality in the data. Minority borrowers do not seem more likely to submit applications toward the end of the month. Nor do we find evidence that omitted variables—including application quality—change within the month. For example, the share of applicants with an income lower than the county median is stable over the course of the month. This income test serves as a proxy for other potential differences between applicants. Furthermore, the ex post default rate gap doesn't vary over the course of the month. Our findings suggest that the shrinking approval rate gap is likely not caused by application- or applicant-related factors.[18] Therefore, using the high-frequency data allows us to attribute the decline in discrimination to loan officers rushing to meet their monthly quotas.

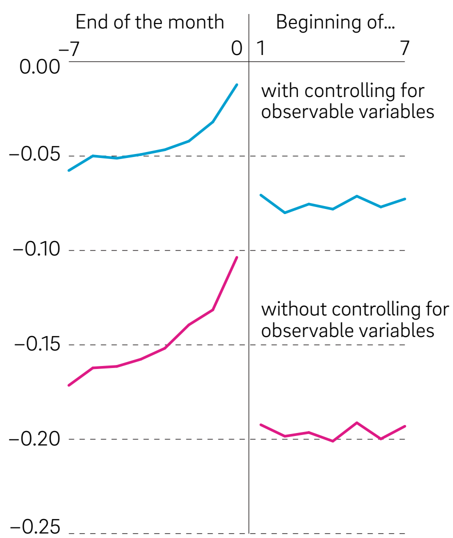
We find that the time-varying discrimination explains about 3.5-5 percentage points of the approval rate gap, which is about half of the unexplained approval rate gap of 7 percent after controlling for observable loan-level characteristics.

Our research also enables us to test the theory that AU reduces discrimination in the mortgage markets. We find that the gap in AU recommendations is nearly constant over the course of the month, which suggests less racial bias in AU decisions. Nonetheless, the approval rate gap of human-made decisions decreases for lenders that use an AU system (though not as much as for lenders that do not).[19] This implies that there can still be racial bias when a human is making the approval decision, even after receiving an AU recommendation (Figure 4). Consistent with previous studies, our research shows that AU seems to reduce but does not eliminate the racial gap in approval rates in the mortgage markets. One possible criticism of our methodology is that the within-month variation in the approval gap merely reflects differences in how long it takes to complete the origination process. For example, the shrinking approval rate gap can be a result of African American borrowers being more likely to settle their housing transactions (and hence mortgage applications) at the end of a month. But we find little evidence that racial differences in the time between application and origination vary within the month.

FIGURE 3

The Approval Rate Gap Shrinks to Nearly Zero When We Control for Other Variables

Approval rate for African American applicants minus the rate for white applicants in the seven days preceding and succeeding the first of the month, controlling for several observable variables, 1994–2019



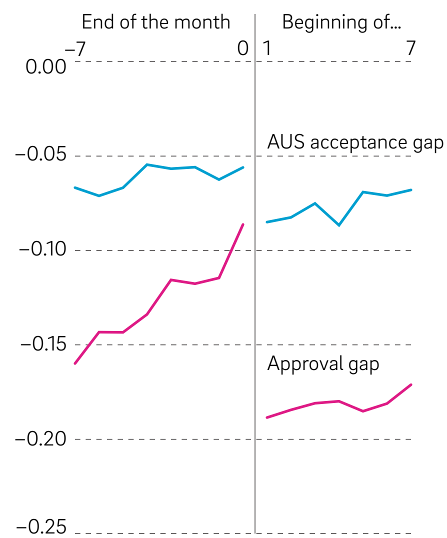
Source: Home Mortgage Disclosure Act (HMDA) data set, Board of Governors of the Federal Reserve System.

Note: Figure 3 plots the average approval rate residual gap from a regression of approval rates on loans and application characteristics from HMDA, such as loan amount, applicant income, and whether the loan was for purchase or refinance.

FIGURE 4

Even When Loan Officers Use AU, They Tend to Reject African American Applicants More Often Than AU Recommends

AU's approval rate gap between African American and white applicants versus the actual approval rate gap among lenders who use AU, 2018–2019



Source: Home Mortgage Disclosure Act (HMDA) data set, Board of Governors of the Federal Reserve System.

Conclusion

Researchers have long documented racial discrimination in the mortgage markets, and that literature is growing as AU prompts them to study its impact on antidiscrimination efforts in the mortgage-lending business. In this article, I summarize work by some earlier and more recent researchers who studied racial discrimination in the mortgage markets. Except for Bhutta and Hizmo's 2021 article, most papers find that there is at least some racial bias in the mortgage markets.

Papers in the literature attempt to reduce the problem of omitted variables largely by adding more control variables to the analysis. However, it is difficult to know whether the additional variables eliminate the bias. My research shows an approach that could solve the problem of omitted variables by using high-frequency data.

A 2021 working paper by Bhutta, Hizmo, and Federal Reserve economist Daniel Ringo, as well as my research using the high-frequency data approach, both show that AU seems to reduce but not eliminate the racial approval rate gap in the mortgage markets. Based on these findings, policymakers might want to encourage the use of AU to help reduce racial discrimination. However, data are available only for the last few years, so research in this area is still relatively new. Further research is needed to confirm our findings.

Notes

1 Some analysts argue that such regulations are unnecessary and that market competitive pressures undermine the desire to discriminate. See, for example, Becker (1957).

2 See McCabe (2019).

3 See Buchak et al. (2018) and Fuster et al. (2019) for examples.

4 Fintech lenders use innovative technology designed to outperform traditional financial methods in the delivery of financial services.

5 According to the law, both "taste-based" and "statistical" discrimination are illegal. In economics, taste-based discrimination refers to discrimination as a result of prejudice, while statistical discrimination refers to decisions that unintentionally correlate with impermissible variables.

6 This would constitute statistical discrimination.

7 See Fuster et al. (forthcoming) for an example.

8 This is in contrast to results from experimental studies, which are more difficult to implement.

9 See Becker (1957).

10 Suppose that there are two, easily distinguishable types of white mortgage applicants: those who have a 1 percent chance of defaulting on a mortgage, and those who have a 50 percent chance. Similarly, assume that African American applicants have either a 5 percent or 50 percent chance of defaulting. If lenders are biased and approve white applicants who have a default rate of no more than 10 percent and African American applicants who have a default rate of no more than 5 percent, these decisions will generate observed ex post average default rates of 1 percent for white borrowers and 5 percent for African American borrowers. This is a case when the ex post default rate is higher for the minority group while there is discrimination, contrary to what the outcome test would suggest if we observe decision thresholds for the marginal borrowers.

11 See, for example, Berkovec et al. (1998).

12 See Ayres (2002) and Canay et al. (2020) for examples.

13 HMDA data are among the earliest and most comprehensive mortgage application data sets in the U.S. One of Congress's goals in enacting the HMDA in 1975 was to identify possible discriminatory lending patterns in the data collected. HMDA data are also used for Community Reinvestment Act bank exams.

14 See Ladd (1998) for a survey of the older literature.

15 See Giacoletti et al. (2021).

16 The figure shows the average loan origination number by calendar days of month. We can also restrict the end of the month to be immediately before the beginning of the month, but the results would look very similar.

17 Similarly, another example that can explain the time-varying approval rate gap is changing underwriting standards. But as shown later in this article, there is no evidence of changes in underwriting standards or application quality within a month.

18 Some regressions in our paper control for additional variables, such as credit scores and low-documentation status, by using a sample that merges HMDA and Black Knight McDash data. The results are similar.

19 The acceptance rate gap of AU shrinks by about 1–2 percentage points within-month, and the approval rate gap decreases by about 6–7 percentage points within-month.

References

Ambrose, Brent, James Conklin, Luis Lopez. "Does Borrower and Broker Race Affect the Cost of Mortgage Credit?" *Review of Financial Studies*, 34:2 (2021), pp. 790–826, https://doi.org/10.1093/rfs/hhaa087.

Ayres, Ian. "Outcome Tests of Racial Disparities in Police Practices," *Justice Research and Policy*, 4:1-2 (2002), pp. 131–142, https://doi.org/10.3818%2FJRP.4.1.2002.131.

Becker, Gary. *The Economics of Discrimination*, 2nd edition. Chicago: The University of Chicago Press, 1957.

Berkovec, James, Glenn Canner, Stuart Gabriel, and Timothy Hannan. "Discrimination, Competition, and Loan Performance in FHA Mortgage Lending," *Review of Economics and Statistics*, 80:2 (1998), pp. 241-250, https://doi.org/10.1162/003465398557483.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace. "Consumer-Lending Discrimination in the Fintech Era," National Bureau of Economic Research Working Paper 25943 (2019), https://doi.org/10.3386/w25943.

Bhutta, Neil, and Aurel Hizmo. "Do Minorities Pay More for Mortgages?" *Review of Financial Studies*, 34:2 (2021), pp. 763–789, https://doi.org/10.1093/rfs/hhaa047.

Bhutta, Neil, Aurel Hizmo, and Daniel Ringo. "How Much Does Racial Bias Affect Mortgage Lending? Evidence from Human and Algorithmic Credit Decisions," working paper (2021).

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru. "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks," *Journal of Financial Economics*, 130:3 (2018), pp. 453–483, https://doi.org/10.1016/j.jfineco.2018.03.011.

Canay, Ivan, Magne Mogstad, and Jack Mountjoy. "On the Use of Outcome Tests for Detecting Bias in Decision Making," National Bureau of Economic Research Working Paper 27802 (2020), https://doi.org/10.3386/w27802.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery. "The Role of Technology in Mortgage Lending," *Review of Financial Studies*, 32:5 (2019), pp. 1854–1899, https://doi.org/10.1093/rfs/hhz018.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther. "Predictably Unequal? The Effects of Machine Learning on Credit Markets," *Journal of Finance*, forthcoming.

Giacoletti, Marco, Rawley Heimer, and Edison Yu. "Using High-Frequency Evaluations to Estimate Discrimination: Evidence from Mortgage Loan Officers," Federal Reserve Bank of Philadelphia Working Paper 21-04/R (2021), https://doi.org/10.21799/frbp.wp.2021.04.

Ladd, Helen. "Evidence on Discrimination in Mortgage Lending," *Journal of Economic Perspectives*, 12:2 (1998), pp. 41–62, https://doi.org/10.1257/jep.12.2.41.

McCabe, Caitlin. "Wells Fargo to Pay Philly $10 million to Resolve Lawsuit Alleging Lending Discrimination Against Minorities," *Philadelphia Inquirer*, December 16, 2019.

Munnell, Alicia, Geoffrey Tootell, Lynn Browne, and James McEneaney. "Mortgage Lending in Boston: Interpreting HMDA Data," *American Economic Review*, 86:1 (1996), pp. 25–53.