Mortgage Lending Discrimination

A Review of Existing Evidence

Margery Austin Turner, Felicity Skidmore

To view the U.S. Department of Housing and Urban Development Summary of the report, please visit their website.

The nonpartisan Urban Institute publishes studies, reports, and books on timely topics worthy of public consideration. The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders; they do not necessarily reflect the official policies or positions of the U.S. Department of Housing and Urban Development or organizations with whom the authors collaborated on this project.

CONTENTS

Chapter 1
Introduction, Summary, and Recommendations
Margery Austin Turner and Felicity Skidmore

Why Is It Difficult to Measure Lending Discrimination?
Different Forms of Discrimination
Possible Reasons for Discrimination
Potential for Discrimination throughout the Mortgage Lending Process

Summary of Evidence
Advertising and Outreach
Pre-Application
The Loan Approval Denial Decision
Loan Administration

Recommendations for Expanding the Knowledge Base
Launch Systematic Research on Office Locations, Outreach, and Referrals
Expand and Refine Paired Testing of Lenders
Conduct a Rigorous Statistical Analysis of Mortgage Approvals Nationwide
Design and Conduct Research on Loan Terms and Conditions
Further Evaluate "Best Practices" for Remediying Discrimination

Chapter 2
New Evidence from Lender Testing: Discrimination at the Pre-Application Stage
Robin Smith and Michelle DeLair

The NFHA Tests
The Urban Institute Reanalysis Data Set
Results
Regional Differences
Lessons for Future Testing
Annex A: Mean Differences between White and Black Testers in Contact Length
Annex B: Information on Requested Loan Amounts by City
Annex C: Reanalysis Data Collection Form

Chapter 3
Does Discrimination in Mortgage Lending Exist? The Boston Fed Study and Its Critics
Stephen L. Ross and John Yinger

Background
Omitted Variables
Evaluation
Reanalysis
Interpreting the "Meets Guidelines" Results

Data Errors in the Explanatory Variables
Evaluation
Reanalysis
Conclusions

Misclassification in the Dependent Variable
Evaluation
Reanalysis
Conclusions

Incorrect Specification
Evaluation: Reverse Regression
Reanalysis: Reverse Regression
Chapter 4

Other Evidence of Discrimination: Recent Studies of Redlining and of Discrimination in Loan Approval and Loan Terms

Stephen L. Ross and John Yinger

Loan Denial Studies Based on HMDA Data
Loan Denial Studies Base on Applications to Individual Lenders
Studies of Discrimination in Loan Terms
Recent Studies of Redlining
  Recent Studies of Process-Based Redlining
  Recent Studies of Outcome-Based Redlining
Conclusions

Studies of the Causes of Discrimination
Conclusions

Chapter 5

The Default Approach to Studying Mortgage Discrimination: A Rebuttal

Stephen L. Ross and John Yinger

Methodological Issues in the Default Approach to Studying Discrimination
  Unobserved Borrower Characteristics
  Sample-Selection Bias and the Power of The Default Approach to Detect Discrimination
  Conventional versus Government-Insured Loans

New Twist on the Default Approach
  Simulations
    A Default Test Based on Market Concentration

Conclusions

Technical Appendix to Chapters 3 through 5

Stephen L. Ross and John Yinger

Econometric Concepts
  Chapter 3: Omitted Variables, Reanalysis
    A Simple Model of Endogeneity in the Loan Denial Equation
    A Complex Model of Endogeneity in "Meets Guidelines"
  Chapter 3: Data Errors in the Explanatory Variables, Reanalysis
  Chapter 3: Endogenous Explanatory Variable, Reanalysis
  Chapter 5: New Twists, a Default Test Based on Market Concentration

Chapter 6

Inside a Lender: A Case Study of the Mortgage Application Process

Kenneth Temkin, Diane K. Levy, and David Levine

Description of Case Study Lender
Lender's Origination Process
  Overview
  Referral
  Initial Application Interview
  Underwriting
  Post-Underwriting
Conclusions

The Lender's HMDA Performance
  Overall HMDA Performance
  HMDA Performance Controlling for Applicant Income
  HMDA Analysis for Government Loans
Conclusions

What the Lender Isn't Doing and Should Do
  Develop a Formal Mission Statement
  Monitor Fair Lending Performance
  Compensate Staff in a Manner Consistent with Fair Lending Objectives
  Develop a Diverse Workforce
  Provide Training in Fair Lending
  Conduct Outreach to Minorities
  Self-Monitor Fair Lending Performance
  Designate an Employee to Receive Fair Lending Complaints
  Implementation Strategies
  Implementation Steps
  Policy Research Conclusions
INTRODUCTION, SUMMARY, AND RECOMMENDATIONS

A major element of the American dream is a home of one's own in the neighborhood of one's choice. Owning a home is one of the primary ways of accumulating wealth in our society, a form of wealth acquisition that is especially protected in the U.S. tax code. Being a homeowner increases people's feelings of control over their lives and their sense of overall well-being. High rates of homeownership are believed to strengthen neighborhoods as well, by increasing residents' stake in the future of their communities.

Not all Americans, however, enjoy equal access to the benefits of homeownership. Federal law prohibits discrimination in the homebuying process, mandating that all would-be homebuyers must be treated equally by real estate agents, lenders, appraisers, and insurance brokers. However, existing enforcement mechanisms may not be effective enough to guarantee equal treatment or equitable results. Indeed, research clearly shows that minorities still face substantial discrimination in the process of looking for a home to buy (or rent).

Many people believe that minorities also face discrimination when they try to obtain a mortgage—a necessity for most Americans wanting to buy a home. There is no question that minorities are less likely than whites to obtain mortgage financing and that, if successful, they receive less generous loan amounts and terms. But whether these differences are the result of discrimination—rather than the inevitable result of objectively lower creditworthiness—is the subject of a raging debate. The problem is that analysts or practitioners have ignored the question of discrimination in mortgage lending. Many research and investigative studies have addressed certain facets of it, using different data sets and analytic techniques to study various outcomes. The problem is that these studies have not produced a clear consensus on a set of conclusions.

The purpose of this volume is to sort through the research evidence on mortgage lending discrimination, in order to provide policymakers with a comprehensive and comprehensible review of the current state of knowledge on lending discrimination and identify important questions that still need to be answered in order to recommend how best to further the goal of fair housing for all.

Our review of the existing research evidence concludes that many institutions in American society in the past have discriminated against minorities that still exist today. This legacy includes racial and ethnic differences in characteristics that influence the creditworthiness of mortgage applicants—income, accumulated wealth, property values in minority neighborhoods, and credit history. Much of the current debate about mortgage lending discrimination stems from disagreement about the extent to which differential success in obtaining a mortgage is due to credit-related factors that vary with race or ethnicity and how much is due to ongoing discrimination.

WHY IT IS DIFFICULT TO MEASURE LENDING DISCRIMINATION

Investigative activities by fair housing advocates and others have identified and successfully prosecuted individual cases of mortgage lending discrimination. However, analytic studies measuring the overall incidence of discrimination are subject to widely differing interpretations. The crux of the problem is that legal evidence of discrimination in specific cases is not the same as statistical measures of the overall level at which discrimination occurs. For analytic estimates of discrimination, researchers need to be confident that individual instances of discrimination are more than isolated occurrences, and that they add up to a consistent pattern that favors whites and outweighs in a statistical sense any corresponding pattern that favors minorities.

Two characteristics of mortgage financing make it especially difficult to reach definitive statistical estimates of discrimination. The first is that the home mortgage lending process is a complex series of stages. Discrimination could be occurring at any one or more of these, and it could take different forms at different stages. But until the stages themselves are clearly distinguished, and the incidence of discrimination measured at each, its overall incidence cannot be properly interpreted. The second is that deliberate discrimination by many institutions in American society in the past has left a legacy of economic inequality between whites and minorities that still exists today. This legacy includes racial and ethnic differences in characteristics that influence the creditworthiness of any mortgage applicant—income, accumulated wealth, property values in minority neighborhoods, and credit history. Much of the current debate about mortgage lending discrimination stems from disagreement about the extent to which differential success in obtaining a mortgage is due to credit-related factors that vary with race or ethnicity and how much is due to ongoing discrimination.

Different Forms of Discrimination

Discrimination in mortgage lending can take two different forms. It is important to understand the distinctions clearly, because the different forms of discrimination may require different measurement strategies, as well as different remedies. The fundamental distinction is between differential treatment and disparate impact discrimination.
Home mortgage lending is a complex process, composed of many different decision points and institutional policies. The potential for discrimination exists at any one or more points along the way. This makes research challenging, for several reasons. First, a finding of little or no discrimination at one stage in the process does not necessarily prove the absence of discrimination in the process as a whole. Moreover, discrimination may take different forms from one stage to the next, so that a single set of measurement techniques may not apply across the entire process. Finally, discrimination at one stage may influence the characteristics and requests of potential borrowers at a subsequent stage. For example, if a lender systematically steers minorities to apply for federally insured loans, while whites are encouraged to apply for conventional loans, analysis of the loan approval decision will be complicated by the fact that minorities and whites are requesting different types of loans, regardless of their qualifications.

**Summary of the Evidence**

Although the loan approval/denial decision is what comes to mind when most people think about the mortgage lending process, the process starts considerably earlier than that, with the preliminary stages filtering out some would-be mortgage applicants before they even get to a loan officer (see exhibit 1 for an overview of key stages in the process).
The process actually begins with advertising and other outreach efforts—how potential mortgage applicants find out about lending institutions and loan alternatives. To some extent lenders use traditional means to advertise loans, such as newspapers and television, which are available on an equal basis to all who care to look. But they may also make special efforts to reach (or avoid) particular segments of the population.

The second stage encompasses the information and encouragement people receive when they call or visit a lender's office to inquire about mortgage loan terms and conditions. Do minorities and whites receive different levels of services and assistance? Are they given different amounts or types of information? Are they told they may qualify for different types of loans? Do they receive different degrees of encouragement and help in understanding how to overcome barriers to applications and approval?

The third stage in the process is the loan approval/denial decision. This stage involves submitting an application that includes a range of information required to determine the applicant's creditworthiness, confirmation (or not) of that information, the lender's up-or-down decision, and, if the loan is approved, which loan product it is.

The final stage is loan administration. Even after a mortgage has been approved and issued, lenders can exercise considerable discretion about how to treat people who have missed one or more payments. They can accept penalties for several months, or they can start foreclosure proceedings.

Here we summarize briefly what is known about discrimination at each stage of the process, including new findings presented in this volume.

Advertising and Outreach

There is compelling legal evidence of discrimination in the placement of branch offices. This evidence comes from an investigation of the Decatur Federal Savings and Loan Association, which began in 1989 with a U.S. Justice Department investigation and ended with a consent decree signed by the two litigating parties in 1992. The investigation found that Decatur Federal had opened 43 branches in the Atlanta metropolitan area between its founding in 1927 and the late 1980s, only one of which was in a predominantly black neighborhood. During the same period, it closed two offices—the one originally opened in the predominantly black neighborhood and another one in a neighborhood that had become predominantly black.

Along with this history of branch closings, Decatur Federal explicitly applied different criteria for closing branches in black neighborhoods than in white neighborhoods. In addition, the Justice Department obtained evidence that Decatur Federal had explicitly excluded black census tracts from its market area, even though it was a large-volume lender able to compete throughout the Atlanta metropolitan area. Finally, a former Decatur Federal account executive told investigators that she was specifically instructed by the bank not to solicit loans south of Interstate 20, an area that included many of Atlanta’s black neighborhoods (Ritter 1996; Siskin and Cupingood 1996).

How frequently does discrimination occur at the initial stage in the mortgage lending process? There is no research evidence, as yet, about the incidence of discrimination during the advertising and outreach stage of the mortgage lending process. This is an area where more research is clearly needed, which can build on the insights from the Decatur case as it defines and devises ways of measuring the incidence of these and similar practices across institutions and markets.

Pre-Application Inquiries

Existing knowledge about lender behavior at the pre-application stage comes primarily from paired testing (also known as fair lending audits) undertaken by fair housing advocacy agencies whose mission is to promote fair housing through a variety of channels, including litigation. Testing has been used widely for analytic as well as investigative studies of discrimination by landlords and real estate agents; however, only a few relatively small-scale investigative studies—primarily by the National Fair Housing Alliance (NFHA)—have been applied to mortgage lending.

The NFHA audits, funded by the U.S. Department of Housing and Urban Development’s Fair Housing Initiatives Program (FHIP),
were conducted by fair housing enforcement organizations using testers who posed as first-time homebuyers and refinancers at the pre-application stage. Testers, matched on ratios that relate a household's income and debts to the desired loan amount, visited lenders in person to inquire about the types and terms of loans for which they might qualify. After each visit, testers answered a set of closed-ended questions and wrote extensive narratives about their experiences. NFHA conducted tests in seven cities (Atlanta, Chicago, Dallas, Denver, Detroit, Oakland, and Richmond), with about two-thirds of the tests centered in Chicago and Oakland.

NFHA concluded that lenders often appeared to be less interested in giving information to black customers than to whites; urged black customers, but not whites, to go to another lender; and emphasized to black customers, but not whites, that application procedures would be long and complicated. According to these investigative audits, blacks were also more likely than equally qualified whites to be told that they did not qualify for a mortgage (before they had filed a formal application), and whites were more likely to be "coached" on how best to handle potentially problematic aspects of their credit profile (Smith and Cloud 1996).

Given their purpose, NFHA's tests were not designed to produce statistically valid measures of the incidence of discrimination across lenders or markets. However, NFHA provided the Urban Institute access to data from a large number of the tests it conducted, enabling researchers to construct and analyze a database of statistically tractable information (see chapter 2). It is important to keep in mind that findings from this analysis apply only to the specific sample of lenders tested by NFHA, which were selected in large part because they had already shown signs of potential discriminatory behavior. Nevertheless, these tests provide convincing evidence of significant differential treatment discrimination at the pre-application stage and highlight the need for further testing with a sample that is large and representative enough to allow for statistical estimation.

The most basic measure of service at the pre-application stage is whether a customer is seen by a lender and given information about specific loan products. In four of the five cities in the reanalysis data set, African American testers were more likely to be denied such information than white testers. In four of the five cities, lenders spent more time with white than with minority testers. And in three of the five cities, lenders provided whites with information about more possible loan products. What about the loan products themselves? The information available does not support detailed comparisons of the terms and conditions offered to whites and blacks, but it does indicate which testers were quoted a product with a 30-year term. Comparing the interest rates quoted for these 30-year mortgages reveals that African American testers were more likely to be quoted higher interest rates than their white counterparts in three of the five cities.

One notable feature of these paired test results is their regional variation. Although several treatment variables show the same general pattern across cities, differences between cities are substantial. This is particularly striking because the two cities with the most tests yield opposite results. In Chicago, all of the statistically significant findings were unfavorable to black testers. In Oakland, differences in treatment were rarely significant and, when they were, they often benefited African American testers. This contrast highlights the need to understand better the regional differences in mortgage lending practices and in the incidence and forms of lending discrimination.

The Loan Approval or Denial Decision

The decision about whether to accept or reject a mortgage loan application has been the subject of an impressive amount of sophisticated statistical analysis. The primary information used in these studies is a repository of data compiled as a consequence of the 1975 Home Mortgage Disclosure Act (HMDA). HMDA mandated the annual reporting of information, by mortgage lending institutions with at least $10 million in assets, on the number and dollar amount of both home mortgage and home improvement loans, by census tract or county. Since passage of Section 1211 of the Financial Institutions Reform, Recovery, and Enforcement Act of 1989, HMDA data have also included the race, gender, and income of mortgage loan applicants.

HMDA data are routinely used to compare a lender's denial rates for minority and white loan applicants, as a measure of their loan performance with regard to minorities. But HMDA data alone cannot prove or disprove the existence of lending discrimination, because they do not provide enough information to control for all relevant differences between white and minority borrowers. Even though HMDA data now include borrowers' race and income, they do not include critical information on the wealth and debt levels of loan applicants, their credit histories, the characteristics of properties serving as collateral, the terms of loans for which applications were submitted, or the underwriting criteria used to determine eligibility. Herein lies a good part of the story behind the fierce analytical debate about what can and cannot be said about discrimination in mortgage lending.

The seminal study in the debate over discrimination at the loan approval stage is the so-called Boston Fed Study, undertaken by researchers at the Federal Reserve Bank of Boston, which was initially released in 1992 and published in final form in 1996 (Munnell et al. 1992, 1996). Other recent studies make valuable methodological and substantive contributions, but the lack of a data set comparable to the one collected for the Boston Fed Study casts a shadow over all of this other research and makes the results difficult to interpret.

The Boston Fed Study began with HMDA data for the Boston area and collected 38 additional variables for each application in the sample, covering the whole array of information needed to control for legitimate differences in applicant creditworthiness. That the Boston Fed sponsored the study and gained the cooperation of area lenders suggests that the lending community did not expect the study to find statistically compelling evidence of discrimination. But it did just that—concluding that race was indeed a statistically significant and fairly large influence in lending decisions, even when a mass of detailed information systematically related to the lending decision was controlled for in the statistical analysis.

The findings of the Boston Fed Study had an explosive effect on the mortgage lending discrimination debate, initially stimulating extensive soul searching by the industry, followed by a great deal of analytic scrutiny of both the study's data quality and its methodological approaches. The study findings have emerged remarkably intact in the face of most of this scrutiny. But certain complex analytical questions remain that some analysts conclude are enough to undermine the credibility of the original findings. Specifically:

- **Omitted Variables.** Key variables that affect the lending decision, and that are correlated with race or ethnicity, may have been omitted from the Boston Fed analysis. If so, the estimated impact of race on the approval decision may be overstated, because it partly reflects the impact of other, legitimate factors that vary with race and ethnicity.

- **Data Errors.** Mistakes in data entry or data coding may have distorted the Boston Fed analysis, possibly leading to overestimates of the importance of minority status. In addition, some loans in the Boston Fed data set may have been incorrectly classified as approved or disapproved.

- **Incorrect Specification.** When analysts “specify” a predictive equation, they have to make assumptions about how different factors interact to influence the approval or denial decision. If the Boston Fed Study's equations were incorrectly specified, they might again overstate the importance of minority status.

- **Endogenous Explanatory Variables.** Some of the variables in the Boston Fed equations that are used to help explain or predict loan approval may in fact be decided at the same time the loan approval decision is made—or in conjunction with that decision. For example, many loan terms, such as the loan-to-value ratio, which changes if the applicant changes the down payment, are the result of negotiation or participation in a special loan program. If this is the case, single-equation models cannot disentangle the independent effects of minority status on loan outcomes.

Because of the importance of the Boston Fed Study, this volume presents a comprehensive review and reanalysis to assess these
The differences in loan denial rates between minority and white applicants found by the Boston Fed Study cannot be explained away by data errors, omitted variables, or interrelationships between factors that influence loan approval (endogeneity).

The Boston Fed Study results do not definitively prove either the presence or the absence of differential treatment discrimination in loan approval, nor do they definitively prove either the presence or the absence of disparate impact discrimination.

BUT, the Boston Fed Study results provide such strong evidence of differential denial rates (other things being equal) that they establish a presumption that discrimination exists, effectively shifting the "burden of proof" to lenders.

If a case as strong as the Boston Fed Study results were made in a courtroom setting, lenders could only escape the conclusion that they were discriminating if they could prove that their actions were based on "business necessity." That is, they would have to prove they used underwriting guidelines with a clear connection to the return on loans, that they applied these guidelines equally to all groups, and that they had no equally profitable guidelines without differences in outcome. In our view, no scholar has come close to showing that the observed intergroup differences in loan approval in Boston can be justified in this way.

The best way to determine whether the observed minority-white differences in loan denial rates are the result of underwriting practices justified by business necessity would be to replicate the Boston Fed Study in other cities, with the addition of loan performance data. This approach would make it possible to determine which observed application characteristics are accurate predictors of loan returns and, therefore, which underwriting guidelines are legitimate. However, such a study still would not be able to distinguish between differential treatment and disparate impact discrimination. The best, and possibly the only, way to isolate differential treatment discrimination in loan approvals is with the paired testing methodology. Specifically, two applicants with the same credit histories and in need of the same type of loan would apply for a mortgage at the same lender. Differential treatment discrimination would exist if minority applicants were systematically treated less favorably in a large sample of cases.

Using Defaults to Measure Discrimination in Loan Approvals. Some researchers have used information on differential default rates as a strategy for determining whether discrimination occurs at the loan approval stage. This approach is premised on the argument that lenders who discriminate against minority applicants do so by effectively raising their underwriting standards—rejecting minorities who meet the standard required of whites and only accepting minorities who meet a higher standard. If this is the case, minorities who receive loans will be less likely to default than whites. Therefore, the argument goes, if minority default rates are the same or higher than those of whites (other things being equal), lenders must be discriminating. Analysis conducted for this project shows that this simple and intuitively appealing argument runs into severe methodological hurdles when used to measure discrimination in mortgage lending (see chapter 5). The difficulty of obtaining complete information on factors that influence default risk, as well as the covariance of many such factors with race, means that the default approach probably understates the incidence of discrimination at the loan approval stage.

Negotiating Loan Terms. At the loan approval stage, lenders do not simply decide whether to make a loan. They also set the terms of the loan, including the interest rate, loan fees, maturity, loan-to-value ratio, and loan type (conventional, adjustable rate, FHA, and so on). This is an important issue, because fair housing complaints often involve unfair terms and conditions for loans, and there is reason to believe that the lending industry may be in the process of shifting from "credit rationing"—where customers perceived to be high-risk are denied loans—toward "risk-based pricing"—where these same customers are simply charged a higher price for loans.

With respect to type of loan, several research studies have examined the probability that a borrower will receive an FHA loan instead of a conventional loan. Both research and lender behavior indicate that FHA borrowers are more likely to have relatively low credit qualifications. As a result, a very large pool of potential testers would be required. But the courts have not yet ruled on this matter and any group that pushes paired testing into the loan approval stage of the mortgage process might face high legal bills, if not worse. It might be possible to conduct tests using people's actual credit characteristics, but this approach would be administratively difficult because testers would still have to be matched to have the same credit qualifications. As a result, a very large pool of potential testers would be required.
Launch Systematic Research on Office Locations, Outreach, and Referrals

discrimination, the forms it takes, and the circumstances in which it is most likely to occur. We recommend five key areas where more evidence that discrimination occurs at significant levels. But serious gaps remain in our collective knowledge about the incidence of discrimination in home mortgage lending persists. Although we do not yet have reliable measures of the incidence of minority loans are made with the express intent to foreclose as soon as possible.

The evidence just summarized, which is discussed at length in subsequent chapters of this volume, provides persuasive evidence in institutions as they try to ensure full and fair service to both minority and white customers.

Lending industry experts and fair housing advocates have identified a number of practices and procedures that lenders should implement to reduce the possibility of discrimination against minority applicants. Our case study reveals that the lender we visited has not fully implemented any of these fair lending best practices. Moreover, the research literature on organizational change contains clear lessons about “what it takes” to effectively change behavior within an institution. Thus, the case study illustrates how a lending institution might be discriminating against minorities despite its best intentions, and it reflects the challenges confronting lending institutions as they try to ensure full and fair service to both minority and white customers.

RECOMMENDATIONS FOR EXPANDING THE KNOWLEDGE BASE

The evidence just summarized, which is discussed at length in subsequent chapters of this volume, provides persuasive evidence that discrimination in home mortgage lending persists. Although we do not have reliable measures of the incidence of discrimination at each stage in the lending process, systematic monitoring and enforcement efforts are clearly justified by existing evidence that discrimination occurs at significant levels. But serious gaps remain in our collective knowledge about the incidence of discrimination, the forms it takes, and the circumstances in which it is most likely to occur. We recommend five key areas where more information and analysis can and should be assembled to inform both public policy and private action.

The Loan Approval/Denial Process from a Lender’s Perspective

It is intriguing that the Boston-area mortgage lenders apparently believed that discrimination would not be found in the investigations of their practices. If they had, it is unlikely that they would have cooperated so fully with the Boston Fed survey. But the evidence reviewed here strongly suggests that their belief that they were not discriminating was false. Is it possible that lenders discriminate unknowingly? Can discrimination occur in the mortgage lending process even when people believe they are treating all applicants fairly? The answer to this question is vitally important in the quest for strategies to eliminate discrimination in home mortgage lending.

In an effort to shed new light on the issue, this project paid a field visit to a mortgage lending institution (see chapter 6). We conducted in-depth, structured interviews about the mortgage lending process to determine what role employees played in decisionmaking, whether they were aware of fair lending requirements, how they perceived fair lending issues, and how they were monitored by their company for fair lending compliance. After the visit, the impressions of our site visit team were compared with standard HMDA indicators of the lender’s fair housing performance.

The loan we visited is a mortgage company, fully owned by a builder who develops housing for low- and moderate- as well as middle- and upper-income households. The lending institution has 31 employees and currently originates mortgages worth about $70 million a year. Its loans are almost all for home purchases rather than refinancing, and it processes more minority applications than the average for its metropolitan area. The loan origination staff includes six loan counselors, who meet with prospective customers and take applications, and four loan processors, who collect the documentation needed to complete the applications. The branch manager of the company supervises both these groups. The company also has an underwriter who is responsible for assessing completed applications for government-insured loans (conventional loans are underwritten by an outside firm). The branch manager and the underwriter both report directly to the company’s president.

Over the course of a two-day site visit, the research team scrutinized the process used to assess applications and was favorably impressed by the combination of a highly transparent review process, a strong commitment to qualifying marginal applicants, and the genuine belief by all staff that their process is color blind. The team’s strong expectation was that the lender’s HMDA data would show a relatively small denial disparity between white and minority applicants. However, that did not turn out to be the case.

The lender’s denial rate for minorities is lower than average for its metropolitan area, indicating that it does a good job of qualifying marginal minority applicants (and/or attracts minority applicants with above-average qualifications). But disparities between its denial rates for whites and for minorities are high, compared to metro-area averages. How can we reconcile these disparities with the lender’s strong belief that its loan origination process contains absolutely no discriminatory treatment of minority borrowers? There are three possible explanations:

- A large share of the lender’s minority loan applicants may actually be poor credit risks. It is possible that because the case study lender serves more minority customers than other area lenders, these customers may be less creditworthy—on average—than minority loan applicants in the metro area as a whole. If so, the case study lender’s high denial disparities (relative to metro-wide averages) may reflect the diversity of its customer base rather than the possibility of discrimination. However, this explanation seems inconsistent with the evidence that the case study lender approves a larger share of applications (from both minorities and whites) than the average for mortgage lenders metrowide.

- The case study lender may be applying underwriting standards that have a disparate impact on minority borrowers. In other words, minority customers may be denied at relatively high rates because some of the underwriting standards applied by the case study lender have a disproportionate effect on minorities and do not serve a clear business necessity. This explanation seems inconsistent with the fact that denial disparities between whites and minorities are significantly lower among other lenders in the metropolitan area.

- The lender’s staff may be providing preferential treatment to white customers without realizing it. Our case study indicates that loan counselors work hard with customers to overcome problems in their applications. It is possible that the counselors are more at ease with white customers than with minorities, find it easier to communicate and sympathize, or feel more comfortable spending time with whites to solve credit problems. If this is the case, then minorities would be at a disadvantage, not because they were treated badly but because whites were treated better.

Given the information currently available, it is impossible to determine with certainty which of these explanations is correct. It is clear, however, that despite the commitment and good intentions of the case study lender, denial rates for minority loan applicants are unusually high, relative to denial rates for white customers. And these denial disparities appear to be out of line with comparable ratios for the metropolitan market as a whole.

Lending industry experts and fair housing advocates have identified a number of practices and procedures that lenders should implement to reduce the possibility of discrimination against minority applicants. Our case study reveals that the lender we visited has not fully implemented any of these fair lending best practices. Moreover, the research literature on organizational change contains clear lessons about “what it takes” to effectively change behavior within an institution. Thus, the case study illustrates how a lending institution might be discriminating against minorities despite its best intentions, and it reflects the challenges confronting lending institutions as they try to ensure full and fair service to both minority and white customers.
Paired testing can and should be expanded at the mortgage pre-application stage. The testing conducted by NFHA demonstrates that paired testing is feasible and that it uncovers instances of differential treatment that might otherwise go undetected. Because at least some lenders provide more information and assistance to white borrowers, minorities may be discouraged from submitting applications or may apply for loans with unfavorable terms. Discrimination at this stage cannot be detected through analysis of HMMA data or data drawn from lenders’ application files. In fact, paired testing may be the only strategy for uncovering the incidence of discrimination at the pre-application stage. NFHA’s testing (and our reanalysis of the NFHA test results) represents an important first step. But more work is needed to refine testing procedures and apply them to representative samples of lending institutions.

Paired testing can be effective for both research and enforcement purposes, although the procedures used for these two purposes are not identical. Research testing is designed to yield statistically reliable measures of the incidence and severity of differential treatment across a large number of transactions. Because all of the lender testing conducted to date was designed primarily for enforcement purposes, there are limits to what it can tell us in this regard. In order to learn more, the federal government should sponsor a paired testing effort whose primary goal is to quantify the incidence and severity of discrimination at the pre-application stage. Indeed, HUD is currently funding a pilot study that will develop several alternative paired testing methodologies and estimate levels of differential treatment at the pre-application stage for at least one market area.

Ultimately, such testing studies should be conducted in multiple markets, so that they can capture variation in levels and patterns of discrimination across sites. As discussed earlier, analysis of the NFHA test results suggests that there may be substantial differences between cities, and these differences need to be investigated more thoroughly. In addition, the lending institutions where tests are conducted should be selected systematically, to be representative of all lenders of a particular type or those serving a particular market. For example, tests might be conducted for a random sample of metropolitan area, for a sample of institutions over a certain size, or for a sample of those reporting a certain number of mortgage loans.

Test reporting forms should be as tightly structured as possible to permit objective comparisons of the treatment received by whites and minorities across a large number of tests. This may require advance research—or “scouting”—on the products offered and procedures followed by lending institutions in the study sites. Unless researchers and test supervisors know in advance how lending institutions treat potential borrowers prior to the formal application stage, what different loan products are called, and to whom potential borrowers might be referred, it is difficult for pairs of testers to make identical requests and to record accurately the treatment they receive. Moreover, testers should receive careful training and supervision to ensure that both members of each pair present the same attributes, qualifications, and financing needs and that both record their treatment fully and accurately.

Finally, more thought needs to be given to the specifics of lender testing scenarios. No single test pair can explore all possible requests that potential borrowers might make at the pre-application stage or all types of lending institutions in the market. The NFHA tests paired minorities and whites posing as relatively uninformed customers who were well qualified for the types of financing about which they were inquiring. This scenario makes sense because it gives lenders the discretion to suggest different products, request different levels of information, or offer different amounts of capture (and possibly different levels) of discrimination. For example, there is good reason to believe that marginally qualified whites receive more assistance and encouragement in correcting credit problems than do marginally qualified minorities. Thus, a study in which partners posed as marginally or poorly qualified borrowers might elicit different responses from lenders than a study in which testers pose as well-qualified applicants. The results of research testing could prove to be extremely sensitive to the specifics of the test scenario.

At the same time that work on research testing proceeds, fair lending enforcement testing should be refined and expanded. Pre-application testing is essential for finding out if lenders are discouraging minority borrowers from even applying, steering minorities to apply for particular loan products, or referring them to other types of lending institutions. Thus, this type of paired testing plays a critical role in the federal government’s efforts to monitor fair lending compliance and to investigate complaints of discrimination. Fair housing organizations should be encouraged and supported in their efforts to conduct rigorous pre-application testing, both in response to complaints and to assess the extent to which differential treatment may be going undetected in the communities they serve. Moreover, lenders should be encouraged to conduct “self-testing,” as a way to monitor the performance of their own lending procedures.

Testing should not be ruled out as a strategy for investigating and measuring discrimination beyond the pre-application stage. As discussed earlier in this report, paired testing appears to be the only research methodology that would disentangle differential treatment discrimination from disparate impact discrimination at the loan approval stage. Federal law makes it illegal to provide false information on a credit application, and many people believe that this precludes full application testing of mortgage lending institutions. However, some testing advocates argue that submitting false information as part of a paired test—when the tester does not actually intend to borrow money or incur any other financial obligation—does not violate this law. So it is possible that some organizations may be willing to incur the risk of conducting paired testing beyond the pre-application stage—or that the federal government could issue guidance that would allow and encourage greater use of testing. Moreover, it may be feasible to design a paired testing study using the actual income and credit characteristics of testers, although the challenge involved in recruiting equally matched testers would be substantial.

Some researchers also have argued for the use of nonpaired testing of mortgage lending decisions. This would involve finding a pool of actual candidates for mortgage loans. The applicants would then file genuine loan applications, and the progress that they made through the loan application and approval process would be monitored and documented. Analysis would focus on differential treatment of applicants from different racial and ethnic backgrounds in loan approvals and, in the case of approved loans, in the loan amount, interest rates, maturity, loan type, and collateral. Nonpaired testing could provide definitive estimates of the overall incidence of discrimination in loan approvals, but only paired testing can reliably distinguish differential treatment discrimination from disparate impact discrimination.
Conduct a Rigorous Statistical Analysis of Mortgage Approvals Nationwide

The Boston Fed Study should be replicated for more cities and enhanced to respond to the critical methodological issues discussed in this report. It constitutes the strongest and most complete analysis of discrimination at the loan approval stage. By assembling data on borrower characteristics and credit histories, it enabled researchers to estimate the extent to which minorities are more likely to be denied a mortgage loan, other things being equal. Despite the unprecedented scrutiny and criticism to which this study has been subjected, our reanalysis shows that it clearly disputes claims that blacks and whites receive equal treatment from the lending industry. However, this study is not able to distinguish differential treatment discrimination from disparate impact discrimination. And it cannot completely eliminate the possibility that high denial rates for minorities result from differences in their ability to meet legitimate underwriting criteria—criteria that meet the business necessity test. Moreover, the Boston Fed Study applies to only one urban area at one point in time. Comparable analysis for a representative sample of market areas is needed to assess the persistence of discrimination over time and across markets.

A multisite study of discrimination in loan approvals should build upon the intensive review and criticism generated by the Boston Fed Study. In particular, a national study should invest significant time and attention in the collection and verification of complete and accurate data on borrower characteristics, loan characteristics, property characteristics, and credit history to guard against omitted variables and data errors that may bias results. Because of widespread differences between whites and minorities in income, wealth, property values, and credit histories, analysis that fails to account fully for these factors may seriously overstate the extent of discrimination in mortgage loan approvals. Moreover, future analysis should explore alternative versions of a loan approval model and test extensively for possible interrelationships among explanatory variables to generate unbiased results.

In order to test the hypothesis that high rejection rates for minorities are entirely due to legitimate underwriting criteria, researchers need to assemble and analyze data on loan performance and defaults as well as information on loan applications and originations. As discussed earlier, evidence of higher default rates among minority borrowers than among whites does not prove the absence of discrimination at the loan approval stage. However, analysis of loan defaults does have an important role to play in the analysis of possible disparate impact discrimination. Specifically, underwriting policies and practices that disproportionately affect minorities even when they are even-handedly applied are discriminatory under the law if they do not serve a business necessity. Thus, if an underwriting criterion or requirement systematically disqualifies more minorities than whites, but does not reliably predict future loan performance, it is discriminatory. In fact, even if a criterion did predict future loan performance, it might be considered discriminatory if it could be replaced by an alternative criterion that had less of a disparate impact or by more refined criteria. Data on underwriting criteria and loan terms, borrower and property characteristics, and long-term loan performance all need to be linked to support a definitive analysis of disparate impacts in home mortgage lending.

Finally, statistical analysis of discrimination in the loan approval process should attempt to distinguish discrimination based on the borrower's race or ethnicity from discrimination based on the racial or ethnic composition of the neighborhood in which a property is located. The existing empirical evidence on redlining (discrimination based on neighborhood composition) remains inconclusive. It may prove difficult to disentangle the effects of applicant race and neighborhood race, because most blacks currently live in black neighborhoods while most whites live in white neighborhoods. Nevertheless, the distinction is an important one from a policy perspective.

Design and Conduct Research on Loan Terms and Conditions

To date, relatively little statistical analysis has focused on the potential for discrimination in loan terms and conditions. Fair housing complaints often involve unfair terms and conditions for mortgage loans, and there are some indications that the lending industry is in the process of shifting from credit rationing to risk-based pricing. In other words, lenders may be more likely to charge higher interest rates and/or fees for customers they perceive to be risky, rather than denying them financing altogether. Thus, it will be increasingly important to understand how interest rates and fees are determined and to analyze the potential for either differential treatment or disparate impact discrimination in this area.

This issue is closely related to questions about credit scoring. Both risk-based pricing and credit-scoring schemes rely on data about the credit characteristics of borrowers to relate to loan performance. More specifically, these schemes predict—or “score”—the risk associated with a particular borrower, based on past experience. Skeptics of risk-based pricing and credit scoring argue that the experience from which these predictive models are based may not be sufficiently diverse to reflect the favorable performance of loans to minorities and that the variables used in these models may put minorities at an unfair disadvantage. Moreover, none of these schemes has been evaluated by scholars. Thus, rigorous, objective analysis of the relationship between various borrower characteristics and loan performance is critically needed. Otherwise, these schemes may simply institutionalize disparate impact discrimination by imposing rules that put minorities at a disadvantage but that do not serve any business necessity.

In addition, researchers need to investigate systematically the uses of risk-based pricing and credit-scoring schemes, analyzing the relationship between the terms and conditions that lenders use to determine interest rates and fees for mortgage borrowers. This type of research should be used to develop methods for analyzing the potential for either differential treatment or disparate impact discrimination. As several existing studies point out, it is not sufficient simply to compare the final interest rates charged to different groups. Instead, analysis should compare final interest rates to the rates originally quoted when borrowers first inquired. And researchers should attempt to collect and analyze information on various loan fees, again exploring differences between “advertised” and “actual” fees.

Further Evaluate "Best Practices" for Remedyng Discrimination

In order to achieve significant reductions in mortgage lending discrimination, regulatory agencies must do a better job of identifying institutions that are discriminating. But, in addition, both regulators and lenders need to know what it takes to eliminate discriminatory practices. To the extent that discrimination is blatant and intentional, designing corrective remedies may be relatively straightforward. But much of the evidence summarized here suggests that lending institutions may be discriminating without realizing it—through policies and procedures that have a disparate impact on minority borrowers, through subtle differences in the level of encouragement and assistance provided to whites and minorities, or through unexamined assumptions about the types of products and terms for which minorities can qualify. Lending institutions may believe that their practices and decisions have been "color blind," and the institutional changes they need to make to eliminate discrimination may not be obvious.

Fair lending advocates and industry experts have identified a set of strategies that lending institutions should implement in order to comply with anti-discrimination laws. Although these “best practices” appear logical and worthwhile, their effectiveness has not been systematically evaluated. Currently, there is a tendency to identify lending institutions as “high performers” if they are implementing a widely accepted set of best practices, not because they have eliminated or reduced past discrimination. In other words, researchers need to compare fair lending performance for institutions with and without these best practices or for institutions implementing different remedial strategies. The goal of this research is to test the efficacy of various remedies and institutional reforms that lenders implement.

Finally, lending institutions need tools they can use to monitor and assess their own anti-discrimination efforts. The “stick” of litigation or regulatory action obviously creates an important incentive for lenders to care about the potential for discrimination in their policies and procedures. But lenders cannot take action if they do not realize that they are discriminating, and neither regulators nor fair housing groups have sufficient resources to investigate all lending institutions. Self-testing is one strategy lenders can and should implement.
NOTES FOR CHAPTER 1


2. Federal law also prohibits discrimination in housing based on sex, family composition, religion, and disability. This volume focuses on the issue of racial and ethnic discrimination.

3. For a comprehensive discussion of the myriad and complex issues involved in legal and analytic investigations of mortgage lending discrimination, see Goering and Wienk (1996).

4. This report does not address potential discrimination by other important actors—such as real estate brokers, appraisers, insurers, and secondary loan institutions—who are not direct decisionmakers in the mortgage lending decision.

5. We use the term "color blind" in this volume to refer to policies and practices that appear to treat people equally regardless of their race or ethnicity.

6. The Policy Statement on Discrimination in Lending (Federal Register 1994) states that a business necessity must be manifest and may not be hypothetical or speculative. Factors that may be relevant to the justification could include cost and profitability.

7. See 18 U.S.C.S. §1014. Note that this would not bar lenders from conducting self-testing.

8. Although assembling such a database presents significant challenges, federal government regulators have sufficient leverage and resources to obtain the necessary information from lending institutions if they make it a priority.

9. For more information on the data and analysis required to test the business necessity of key underwriting standards, see Temkin et al. 1998.