The Color of Credit

Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement

Stephen Ross and John Yinger

The MIT Press
Cambridge, Massachusetts
London, England

1 Introduction

1.1 Mortgage Discrimination and the American Dream

The last decade has witnessed a flowering of scholarly research on racial and ethnic discrimination in mortgage lending. The findings in this literature are obviously of great importance to millions of American families and also concern community groups, lenders, and public officials. This book reviews, interprets, and extends this literature. Our analysis has profound implications for fair-lending enforcement. We explain some serious flaws in the current enforcement system and show how to fix them.

The topic of discrimination in mortgage lending has, of course, great intrinsic interest. The United States is a nation of homeowners, and the homeownership rate stood at an all-time high of 67.8 percent in 2001 (U.S. Census Bureau, 2001, table 20). Homeownership is the most commonly used method for wealth accumulation,[1] and it is widely viewed as critical for access to the nicest communities and the best local public services, especially education.[2] In addition, one study (Green and White, 1997) finds evidence that homeownership provides nonfinancial benefits, such as problem-solving and management skills, that help the children of homeowners stay in school and stay out of trouble. Another recent study (DiPasquale and Glaeser, 1999) presents evidence that homeownership is associated with investments in social capital, such as membership in nonprofit organizations or helping to solve local problems.

Homeownership and mortgage lending are linked, of course, because the vast majority of home purchases are made with the help of a mortgage loan. In fact, about 8.3 million applications for home purchase mortgages were received by lenders in 2000 alone (FFIEC, 2001b, tables 4-1 and 4-2).[3] In this setting, barriers to obtaining a

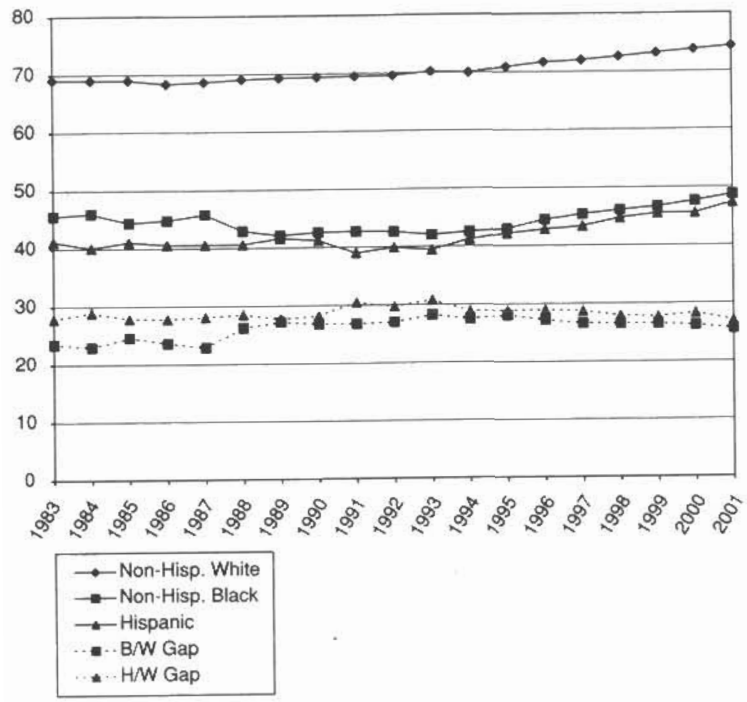

Figure 1.1
Homeownership rates, 1983–2001.
Source: HUD (2002).

mortgage must be taken seriously because they represent obstacles to attaining the American dream of owning one's own home. These barriers take on added urgency when they are related to race or ethnicity, that is, when they draw on and compound social divisions that have troubled this nation for centuries.

A hint about the potential power of mortgage discrimination is provided by the long-standing gaps in homeowership rates between black and white households and between Hispanic and white households.[4] In 2001, the homeownership rate for non-Hispanic whites, 74.3 percent, was 26.6 percentage points higher than the homeownership rate for blacks, 47.7 percent, and 27.0 points higher than the rate for Hispanics, 47.3 percent (U.S. Census Bureau, 2002, table 20). As shown in figure 1.1, these gaps have exceeded twenty-five percentage points since 1988 and are now higher than they were in the early 1980s.[5] They could be caused, of course, by many factors

other than discrimination in mortgage lending, such as intergroup income differences resulting from past discrimination in labor markets or current discrimination by real estate brokers. Nevertheless, the magnitude and persistence of these gaps suggests that all the factors potentially contributing to them, including mortgage discrimination, should be carefully investigated.

1.2 Seeds of Change

Despite the intrinsic importance of discrimination in mortgage lending, few scholars addressed the topic before 1990. Although, as we will see, some important research on the topic was conducted before this time, articles and books on the topic were few and far between, and the results of this research were not widely known.

The first seed for the new burst of attention to the subject was planted in 1988 by a series of Pulitzer Prize–winning articles called "The Color of Money," which were published in the *Atlanta Journal-Constitution*.[6] These articles described practices by Atlanta's lending institutions that appeared to be discriminatory and documented a relatively small flow of mortgage funds to Atlanta's black neighborhoods. Similar articles later appeared in several other newspapers.[7]

A second seed was planted in 1989 when Congress amended the reporting requirements in the 1975 Home Mortgage Disclosure Act (HMDA). The original purpose of this act was to provide information about the geography of loan origins, so it required the vast majority of lenders to indicate the location of each property associated with a mortgage application and the ultimate disposition of the application.[8] This purpose was reendorsed when HMDA was renewed in 1980 and made permanent in 1987. In 1989, the Financial Institutions Reform, Recovery, and Enforcement Act amended HMDA by requiring lenders to provide information on the race and ethnicity of each loan applicant.[9] With these new requirements, therefore, lenders had to report their loan denial rates by racial and ethnic group, as well as by location.

When the HMDA data collected under these amendments were first released in 1991, they revealed a striking disparity in the loan denial rates for different groups. In fact, the denial rate for blacks on conventional home purchase loans was almost two-and-a-half times the rate for whites, and the Hispanic rate was 50 percent higher than the white rate (Canner and Smith, 1991). Although, as we will see, these results do not provide definitive evidence of discrimination, they are still quite troubling, and they received a great deal of publicity at the time they were released. They were, for example, the focus of front-page articles in the *New York Times* and the *Wall Street Journal* (see Quint, 1991, and Thomas, 1991, 1992a, 1992b).

The third seed was planted by the Justice Department. Thanks to the "Color of Money" series, the Justice Department initiated an investigation of Decatur Federal Savings and Loan, one of the Atlanta lenders that appeared to be discriminating, based on the information in the series. The 1990 HMDA data were released in the middle of this investigation, adding motivation—and evidence—to the Justice Department's case. Ultimately, the Justice Department was able to document extensive discrimination, and Decatur Federal signed a consent decree in the fall of 1992: the first high-profile settlement of a mortgage discrimination case.[10] Under this consent decree, Decatur Federal paid $1 million to 48 black applicants whose loans had been denied and altered its marketing practices (Ritter, 1996). This case also made a major contribution to the understanding of discrimination in mortgage lending through the development of new enforcement tools, including regression analysis of loan approval decisions.

The 1990 HMDA data also made an impression on many officials in the federal financial regulatory agencies, including the Federal Reserve Board and the U.S. Department of Housing and Urban Development (HUD). With support and advice from people in these agencies and from some academics, researchers at the Boston Federal Reserve Bank planted a fourth seed. Specifically, they decided to supplement the HMDA data with extensive information on individual loan applications so that it would be possible to isolate discriminatory behavior by lenders.[11] The resulting study, widely known as the Boston Fed Study, first appeared in 1992 (Munnell et al., 1992), and a revised version was published in 1996 (Munnell et al., 1996). The revised version of this study received a great deal of attention in part because it was published in the *American Economic Review*, which is widely regarded as the leading journal in economics.

The original version of this study concluded that "even after controlling for financial, employment, and neighborhood characteristics, black and Hispanic mortgage applicants in the Boston metropolitan area are roughly 60 percent more likely to be turned down than whites" (Munnell et al., 1992, p. 2), a result that can be interpreted as a sign of lending discrimination. After the study was revised in response to the comments of many readers, including anonymous referees for the *American Economic Review*, the final version concluded that black and Hispanic applicants were about 80 percent more likely to be turned down than were comparable whites.

1.3 Profound Change and Unanswered Questions

These seeds have produced profound change in the public and academic debate about mortgage lending discrimination. To cite a few examples: The agencies involved in fair-lending enforcement, including the Federal Reserve, have developed new enforcement procedures and filed fair-lending complaints against several major lenders. Secondary mortgage market institutions have developed programs to promote loans for low-income and minority households. HUD has sponsored research on mortgage discrimination, and dozens, if not hundreds, of articles about mortgage lending discrimination have been published.[12]

For at least two reasons, however, all this activity has failed to produce a consensus on the magnitude or consequences of mortgage lending discrimination, and debate on these subjects remains very lively. First, many commentators believe that trends in the mortgage market since 1990, when the Boston Fed Study's data were collected, have resulted in a reduction in discrimination, but no such decline is apparent in the HMDA data. Second, a huge number of publications, both academic and nonacademic, have commented on the Boston Fed Study's findings and methodology, but opinions on the study vary enormously. Some commentators praise it as a significant advance in the study of lending discrimination. Others criticize its data and methods, and some even conclude that its conclusions are incorrect or at least highly misleading. Moreover, this literature has raised a variety of important new issues, but the data are not yet available to resolve them. This section pursues these two points. We review the HMDA data for the last several years and introduce the recent literature on mortgage lending discrimination, which is the foundation of this book.

1.3.1 *The HMDA Data since 1994*

Despite the attention given to lending discrimination over the last decade by lenders, financial regulators, federal officials, secondary mortgage market institutions, and community groups, mortgage

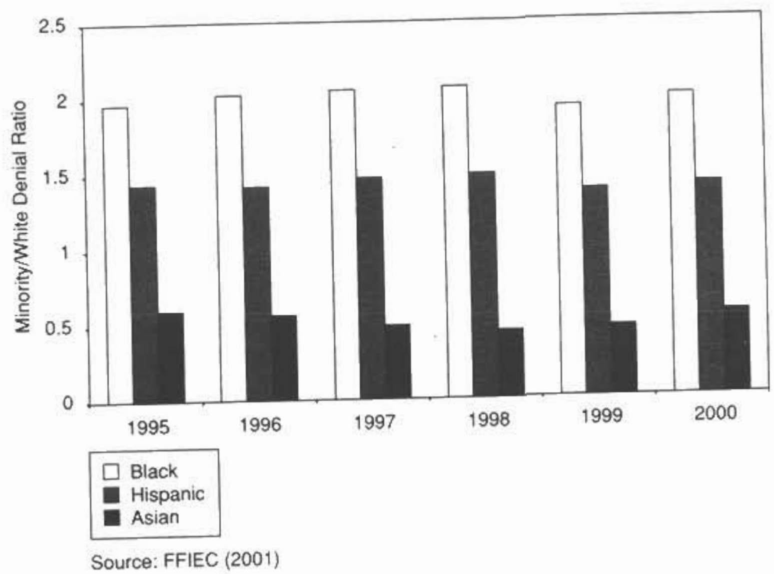

Figure 1.2
Conventional home purchase loan denial ratios by year.
Source: FFIEC (2001b).

loan applications from black and Hispanic households are still much more likely to be denied than are applications from whites. For conventional home purchase loans in 2000, the loan denial rate for blacks divided by the loan denial rate for whites (the black/white denial ratio, for short) was 2.00, which indicates that blacks were twice as likely as whites to be turned down for a loan.[13] Similarly, the Hispanic/white denial ratio was 1.41. In contrast, the Asian/white denial ratio was only 0.55, indicating that applications from Asians were far more likely to be accepted than were applications from whites. These ratios are summarized in figure 1.2.

These numbers cannot be directly compared to the 1990 numbers, because the coverage of the HMDA data was greatly expanded in 1992.[14] Clear comparisons can be made over the last several years, however, and, as shown in figure 1.2, the black/white denial ratio has fluctuated around 2.0 since 1995, with a high of 2.07 in 1998 and a low of 1.92 in 1999. The current ratio, 2.0, is slightly higher than the 1995 ratio, 1.95. The Hispanic/white denial ratio has fluctuated around the lower value of 1.5, but it exhibits a similar pattern over time, with a relatively high value in 1998 and a relatively low value

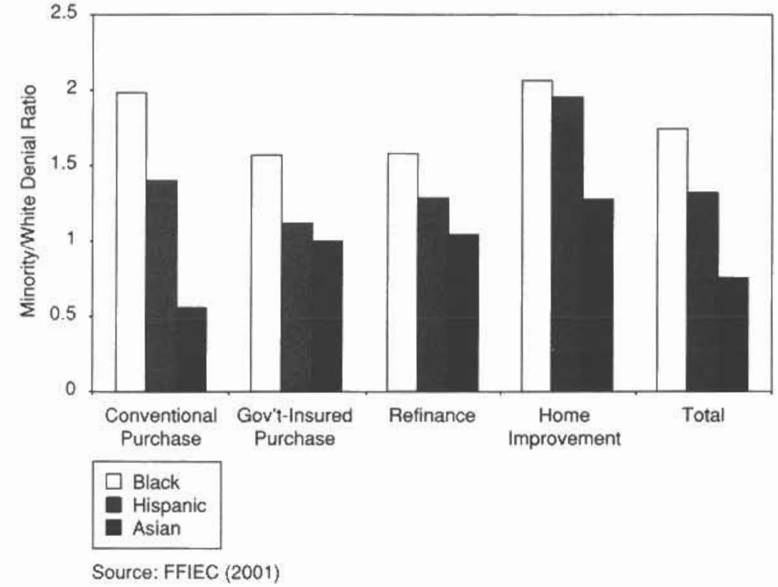

Figure 1.3
Loan denial ratios by loan type, 2000.
Source: FFIEC (2001b).

in 1999. Its current value, 1.41, is slightly below its value in 1995, 1.43. The Asian/white denial ratio has stayed fairly close to 0.5, but it follows a smoother pattern. It declined steadily from 0.61 in 1995 to 0.45 in 1998 and then increased steadily to 0.55 in 2000.

Two other persistent results in the basic HMDA data raise additional questions about mortgage lending that have intrigued scholars and policy makers (Avery, Beeson, and Sniderman, 1996a). First, blacks and Hispanics are also more likely than whites to be turned down for government-insured home purchase loans, refinance loans, and home improvement loans. For these types of loans, Asians also face a higher denial probability than do whites (see figure 1.3). These results raise, but do not answer, two important questions: Is there discrimination in the market for government-insured loans? Do black, Hispanic, and Asian households face discrimination when they apply for refinancing or a home improvement loan, even though these applicants have, by definition, demonstrated their creditworthiness by receiving a home purchase mortgage and then meeting the obligations associated with it?

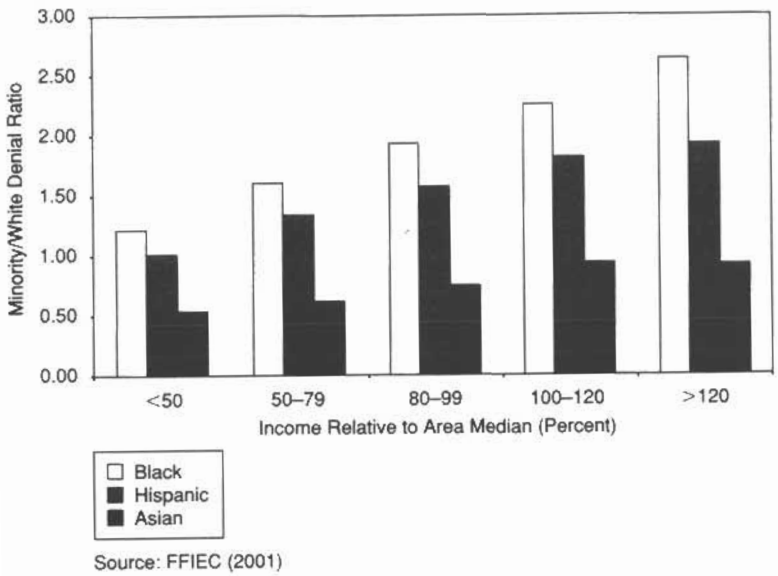

Figure 1.4
Loan denial ratios by income class, 2000.
Source: FFIEC (2001b).

Second, loan denial rates are higher for black and Hispanic applicants than for white applicants at all income levels. Moreover, for blacks, Hispanics, and Asians, the minority/white denial ratio increases steadily with income. This increase is particularly striking for blacks; the denial ratio is only 1.19 in the lowest income category but climbs to 2.48 in the highest category (see figure 1.4). Thus, the higher denial rates for blacks and Hispanics do not appear simply to reflect the fact that these groups have lower average incomes than do whites.[15]

Of course, these higher denial rates also do not prove that blacks and Hispanics face discrimination in mortgage lending, because they do not account for possible differences in loan features or borrower creditworthiness across groups. The differences are so dramatic, however, that they focus attention on the possibility that this type of discrimination might exist. Indeed, the academic literature reviewed in this book is largely devoted to finding out the extent to which the loan approval disparities in the HMDA data and in other indicators of lender behavior reflect racial and ethnic discrimination.

1.3.2 *The Literature since the Boston Fed Study*

For at least three reasons, no consensus has emerged from the huge literature on mortgage lending discrimination that has appeared since the original version of the Boston Fed Study was released.

First, opinions on the Boston Fed Study vary widely, to say the least.[16] One commentator declared, in the title of his piece, that the study "deserves no credit" (Liebouitz, 1993); another argued that the study is "invalid" (Becker, 1993a). In contrast, one review and reanalysis of the Boston Fed Study's data concluded that the study "clearly demonstrated" the existence of discrimination (Carr and Megbolugbe, 1993). To some degree, this range of opinion reflects the inherent complexity of the topic. As we will see, any study of mortgage lending discrimination must grapple with a complex set of methodological issues, and scholars do not agree about which methodology is best for the issues and data involved. They also do not agree on the best way to interpret the Boston Fed Study's findings. Nevertheless, many of the issues in this debate concerning methodology and interpretation can be clarified, if not resolved, by a careful examination of the Boston Fed Study's data, which have been made available to interested scholars.

Second, disagreements about the Boston Fed Study persist because no comparable data set has yet been assembled. A great deal of research on mortgage lending discrimination, some of it very informative, has been conducted since the initial report on the Boston Fed Study was released. Nevertheless, the study has not been replicated for another place or time, let alone improved. All subsequent studies that apply to more than one lender must rely on data that are less complete than the data used for the Boston Fed Study. Given the importance of the topic and the striking nature of the Boston Fed Study's findings, this lack of replication is unfortunate and puzzling. Many institutions, including the federal financial regulatory agencies, the large secondary mortgage market institutions, and HUD clearly have the authority to collect the type of data needed for such a replication. Why haven't any of these institutions collected such data? We do not know the answer to this question, but we do know that scholars and policymakers cannot come to credible conclusions about the current importance of mortgage lending discrimination without access to data of this type.

Third, the debate about mortgage discrimination must grapple with a series of striking changes in mortgage markets over the last decade or two. These changes, which include dramatic growth in nondepository lenders and in automated underwriting, have raised new issues that the Boston Fed Study was not designed to address. Some scholars argue that these changes will lead to less discrimination; others argue that they will lead to more. The impact of these changes on mortgage discrimination has not been widely studied, however, and these arguments have hardly been addressed, let alone resolved, with empirical evidence.

1.4 The Plan of This Book

The main purposes of this book are to explain what has been learned about mortgage lending discrimination in recent years and then, building on that understanding, both to reanalyze existing data and to devise new tests for discrimination in contemporary mortgage markets. Although we do not have any new data to work with, we show that many of the contentious issues that appear in the recent literature can be addressed with the publicly available data collected for the Boston Fed Study, merged, in some cases, with other data that are also publicly available. Much of the recent debate concerns the best methodology for estimating mortgage lending discrimination. We blend the strengths of several existing methods to develop new, straightforward procedures that provide clear estimates of discrimination and avoid the statistical and interpretive problems that have plagued other approaches. These procedures rely on data that could easily be collected by governmental agencies or secondary mortgage market institutions and that could, in principle, be made available to scholars. In addition, these procedures could eliminate a major weakness in the current fair-lending enforcement system, namely, that it looks for only one, narrowly defined type of discrimination. They could also be applied not only to discrimination in loan approval, but also to discrimination in credit-scoring schemes or in loan pricing.

We begin, in chapter 2, by providing some background information on mortgage markets and on the definition of discrimination. We present some basic facts about the operation of the mortgage market and explore recent trends in this market, with a focus on trends that may affect minority households' access to mortgage

loans. These trends are important in part because they are one source of the existing confusion about the importance of mortgage lending discrimination. In addition, we describe in some detail both the laws against lending discrimination and the enforcement system these laws create. The chapter concludes by presenting a preliminary framework for estimating the extent of lending discrimination, employing a definition of discrimination that is consistent with existing fair-lending legislation. This framework is then used to introduce some of the methodological issues that appear throughout the book, such as the biases that can arise when a study does not control for all relevant credit characteristics.

In chapter 3, we lay down a conceptual foundation for our work by reviewing the literature on mortgage markets in general. This review is designed to highlight the complex setting in which a loan approval decision takes place. More specifically, we build on the literature to derive a series of lessons, which we draw on in later chapters, concerning the specification of tests for discrimination in loan approval. For example, we show that a study of lending discrimination may not be able to eliminate biases without controlling not only for variables that directly influence the loan approval decision, but also for variables that influence lender actions during earlier stages of the lending process.

Chapter 4 turns to the early literature on mortgage lending discrimination, defined as the literature up to and including the Boston Fed Study (Munnell et al., 1996). We review all the early studies of which we are aware and discuss the major contributions to the literature. Not surprisingly, many, if not most, of the issues we struggle with in this book were first raised by one or more of these early studies. The chapter also describes the Boston Fed Study, but an evaluation of this study is postponed until the following chapter.

As noted above, the Boston Fed Study has been widely criticized. In chapter 5, we evaluate in detail every criticism of which we are aware, including those concerning problems arising from omitted variables, from data errors in the explanatory variables, from misclassification in the dependent variable, from incorrect specification, and from endogenous explanatory variables. We examine the claims of both the critics and the defenders of the Boston Fed Study. We then formally model the issues raised by each criticism and employ the public-use version of the Boston Fed Study's data set to determine which of these criticisms have merit. Although several other studies have explored potential flaws in the Boston Fed Study, our evaluation is the most comprehensive review yet attempted.

We find that several critics of the Boston Fed Study raise legitimate issues that are worthy of further exploration. Our analysis of these issues leads to two principal conclusions. First, the large minority-white disparity in loan approval found by the Boston Fed Study cannot be explained by data errors, misclassification, omitted variables, or the endogeneity of loan terms. Second, the interpretation of the Boston Fed Study's results depends heavily on an issue not adequately considered by the study's authors, namely, whether different lenders use different underwriting standards. More specifically, the Boston Fed Study cannot rule out the possibility that the minority-white disparity in loan approval reflects variation in underwriting standards across lenders, not discrimination.

The issue of across-lender variation in underwriting standards is so important that we give it a chapter of its own, chapter 6. This issue has been one of the central themes of the literature since the Boston Fed Study, and we begin by reviewing this literature in some detail. We then merge the public-use version of the Boston Fed Study's with the comparable HMDA data, a step that allows us to identify lenders, and provide new estimates of the impact of across-lender variation in underwriting standards on estimated minority-white disparities in loan approval. We find evidence that underwriting standards do, indeed, vary across lenders. We also find, however, using several different methods, that accounting for this variation has no impact on the estimated loan approval disparity. We conclude that the minority-white disparity in the Boston Fed Study's data does provide strong evidence of discrimination in loan approval.

Chapters 7 and 8 complete our review of recent literature. Chapter 7 concentrates on dimensions of discrimination that are related to but distinct from discrimination in loan approval. Specifically, it explores recent literature on redlining, defined as lending discrimination based on a property's location, and on discrimination in loan terms.[17] This literature is in its infancy, but several studies suggest that redlining and discrimination in the setting of mortgage interest rates appear to occur in some circumstances. In addition, this chapter reviews the literature on the causes of discrimination in loan approval. For example, we explore the hypothesis, developed in several recent articles, that discrimination arises because white lend-

ing officials often lack a "cultural affinity" with minority applicants. Clear evidence on the causes of lending discrimination does not yet exist, but our review of the conceptual literature provides background both for the subsequent chapter and for the policy issues discussed at the end of the book.

Chapter 8 addresses an alternative approach to studying discrimination in loan approval, namely, an investigation of minority-white differences in loan defaults. According to this approach, discrimination in loan approval involves holding minority applicants to a higher standard than white applicants. If discrimination exists, the argument goes, minority loans, that is, approved applications, will be of higher quality than white loans, as determined by their default rate. This "default approach" has received a great deal of attention, in part because several scholars have claimed that it refutes the Boston Fed Study's finding that discrimination exists.

The literature on the default approach performs a valuable service by bringing loan performance information into the discussion. Despite its intuitive appeal, however, this approach runs into unsurmountable methodological obstacles. Most importantly, we show that the default approach cannot detect discrimination unless some underwriting variables are excluded from the analysis but yields biased results if the variables that are excluded are correlated with minority status. No study even observes the correlation between excluded variables and minority status, let alone demonstrates that this correlation equals zero. Several recent studies have introduced creative methods for overcoming these obstacles, but we show that all of these methods meet with limited success, at best. We conclude that an analysis of loan defaults alone cannot provide credible evidence concerning the existence of discrimination in loan approval.

Chapter 9 brings together the literatures on loan approval, loan performance, and types of lending discrimination. We first show that a certain type of discrimination can easily be incorporated into a credit-scoring or other automated underwriting system, even one derived from seemingly group-neutral statistical procedures. This type of discrimination is likely to be missed by most current research and enforcement procedures, but it can serve as a substitute for other types of discrimination that these procedures are designed to uncover. Second, we devise new tests for discrimination based on a data set combining information on loan performance, as measured, say, by defaults, and on loan approval. These tests capture all types of discrimination and avoid the methodological problems that have plagued other approaches. Finally, we bring several key themes of the book together by demonstrating that our new tests can provide an accurate estimate of discrimination even if different lenders use different underwriting standards.

In our final chapter, chapter 10, we apply our results to an analysis of the current federal fair-lending enforcement system. We begin by reviewing court decisions and regulations concerning fair lending. Most of the relevant court cases and virtually all of the relevant regulations have been developed for cases involving discrimination in employment. We explore what is known about the application of these cases to discrimination in mortgage lending and develop principles to guide regulations for the enforcement of fair-lending legislation. Moreover, we also draw on our analysis in chapters 6 and 9 to develop fair-lending enforcement procedures that are consistent with existing legal standards. Specifically, we develop two new enforcement tools that capture all types of discrimination if (and only if) they exist. One of these tools is based on loan approval data, and the other combines data on loan approval and on loan performance. We also explain how tools blending approval and performance data could be adapted to consider discrimination in the scores produced by automated underwriting systems and in loan pricing.

Chapter 10 also offers an evaluation of existing fair-lending enforcement procedures. We show that these procedures are seriously inadequate in the sense that they are incapable of identifying many cases of discrimination. Indeed, as currently designed, the enforcement system picks up only extreme cases of a certain type of discrimination in loan approval among a subset of lenders! To put it another way, existing procedures completely miss certain types of discrimination in loan approval and virtually all discrimination in credit scoring and other automated underwriting practices or in loan pricing. We do not know how much discrimination actually exists in these types of behavior, but our analysis in chapter 9 demonstrates that discrimination is easy to introduce into an automated underwriting scheme, even one with apparently group-neutral procedures. It is certainly inappropriate for enforcement officials to ignore large categories of discrimination without any evidence about their incidence. The implementation of our new enforcement tools, or equivalent ones, is required to preserve access to the American dream for all Americans.