http://clevelandfed.org/research/review/
Economic Review 1997 Q4

Beneath the Rhetoric: Clarifying the Debate on Mortgage Lending Discrimination

by Stanley D. Longhofer and Stephen R. Peters

Stanley D. Longhofer is an economist at the Federal Reserve Bank of Cleveland and Stephen R. Peters is a visiting assistant professor of finance at the University of Illinois at Urbana–Champaign.

Introduction

Home ownership has long been identified as a central element of the "American Dream"; in fact, many treat the two as synonymous. No wonder, then, that much of federal antidiscrimination law focuses on housing and the lending markets used to finance home purchases. For more than 30 years, bank regulators have been charged with the task of rooting out illegal discrimination in credit markets.

The long-standing debate over whether lenders systematically discriminate against minorities received added momentum in 1992 with the release of the now-famous "Boston Fed Study" (Munnell et al. [1992]), which purported to find that minority applicants in the Boston area are roughly half again as likely as similarly situated white applicants to be denied mortgage loans. Although their conclusions have been hotly debated in the ensuing years, there can be no question that Munnell et al. put the mortgage discrimination debate on the front burner in both policy and academic circles.

This debate still rages today. Yet after more than six years, we are not much closer to knowing whether discrimination is a widespread problem in mortgage markets. Indeed, economists have yet to agree on whether it is more effective to detect discrimination using denial rate analyses like those employed by Munnell et al. or by examining borrower default rates. Nor have we settled on the appropriate policy response to make if discrimination is in fact as prevalent as Munnell et al. suggest.

A key roadblock in answering these questions has been the lack of a solid theoretical foundation on which to conduct the debate. Without a consistent economic theory of the mortgage application, underwriting, and default processes, fundamental tasks such as defining discrimination and outlining how it might be detected have proven slippery. As a consequence, more advanced issues such as identifying the source of discrimination and evaluating potential policy responses have been obviated, despite longstanding calls to "move beyond a debate over whether discrimination exists to a discussion of how best to eradicate it" (Galster, [1993, p. 146]).

In this article, we outline a simple theoretical model of the mortgage underwriting process, originally developed by Longhofer and Peters (1999), which provides a framework for clearly defining discrimination and various notions of the default rate. By giving those with differing views a common framework for arguing their positions, this model clarifies and reconciles some of the most contentious questions at the heart of the controversy over mortgage discrimination. We also show how this theoretical foundation can aid in designing practical policy responses to this vexing social problem.

The next section reviews the highlights of the debate that ensued in the wake of the initial study by Munnell et al. Section II introduces the Longhofer–Peters model of mortgage underwriting and shows how it can be used to clarify many of the definitions and assumptions that underlie this debate. Section III reexamines the mortgage discrimination debate in light of that model, and section IV concludes.

I. The Mortgage Discrimination Debate

The debate over whether lenders systemically discriminate against minority applicants began in earnest with the release of Munnell et al. (1992).[1] The authors of this study found that minority applicants in the Boston metropolitan area were approximately 50 percent more likely to be rejected than whites, even after controlling for the factors that banks use to determine an applicant's creditworthiness. "This means that 17 percent of black or Hispanic applicants instead of 11 percent would be denied loans, even if they had the same obligation ratios, credit history, loan to value, and property characteristics as white applicants. In short, the results indicate that a serious problem exists in the market for mortgage loans, and lenders, community groups, and regulators must work together to ensure that minorities are treated fairly" (Munnell et al. [1992, p. 44]).

In his Nobel Lecture, Becker (1993b) challenges these findings. Recasting his economic theory of discrimination in the context of the mortgage lending market, Becker argues that bigoted lenders are willing to sacrifice the profit they might earn by approving marginally qualified minority applicants in order to satisfy their "tastes for discrimination." Thus, the proper way to detect discrimination is to do so directly, by comparing the relative profitability of loans to minorities and whites: "This requires examining the default and other payback experiences of loans, the interest rates charged, and so forth. If banks discriminate against minority applicants, they should earn *greater* profits on the loans actually made to them than on those to whites. The reason is that discriminating banks would be willing to accept marginally profitable white applicants" (Becker [1993b, p. 389]).

If banks discriminate against minority borrowers, Becker argues, we should observe minorities defaulting less frequently than whites on average. "[T]he theory of discrimination contains the paradox that the rate of default on loans approved for blacks and Hispanics by discriminatory banks should be lower, not higher, than those on mortgage loans to whites. The reason is that such banks only accept the very best minority candidates" (Becker [1993a]). Because discriminators who are following their tastes will pass up better-qualified minority applicants for less-qualified white ones, the average creditworthiness of the minorities who actually receive loans will be higher than that of whites.

Macey (1994) uses Becker's argument to refute the principal conclusions of Munnell et al., citing that study's own data.[2] "The default rates for white and black mortgage loan applications are equal across census tracts. If bankers were discriminating by turning down marginally qualified black applicants while accepting marginally qualified white applicants, then default rates among whites would be higher. But bankers are accepting the same level of risk from both black and white applicants" (Macey [1994]).

The validity of this argument regarding average default rates relies on the assumption that white and minority credit risk is equally distributed in the borrower pool. In fact, however, Munnell et al.'s data suggest that minority applicants (and borrowers) are less creditworthy than whites on average (see Munnell et al. [1996], table 1).

Galster (1993) uses this fact to argue that minorities will default more often than whites in the absence of discrimination. Thus, he concludes that equal average default rates across minority and white borrower pools imply taste-based discrimination against minorities. On these grounds, Galster suggests that default rates cannot be considered a reliable indicator of discrimination.[3]

The above arguments revolve around whether average default rates can be used to uncover discriminatory behavior on the part of lenders. Others, including Calomiris et al. (1994)

1 Later revised and published as Munnell et al. (1996).

2 See also Brimelow and Spencer (1993) and Becker (1993a).

3 Tootell (1993), Browne and Tootell (1995), and Munnell et al. (1996) echo Galster's conclusions.

and Ferguson and Peters (1995) focus on the interpretation of *marginal* default rates. These authors argue that, regardless of the relative distributions of creditworthiness across racial groups, in the absence of taste-based discrimination the marginal borrower—the least creditworthy applicant to be approved for a loan—will have the same creditworthiness across groups. That is, if banks hold all applicants to the same credit standard, the marginal default rate will be the same for all racial groups. If this is true, then comparisons of the marginal default rate across races may enable regulators to detect taste-based discrimination (in the manner of Becker [1993b]).

In summary, a variety of apparently conflicting tests have been proposed for inferring whether lenders discriminate against minority applicants: Lower average default rates for minorities are evidence of discrimination against minorities (for example, Becker [1993b]); equal average default rates for minorities and whites are evidence of discrimination against minorities (Galster [1993]); and, lower marginal minority default rates are evidence of discrimination against minorities (Calomiris et al. [1994]). Which of these conclusions are we to believe? How should we go about uncovering illegal discrimination? In the next section, we develop a simple theoretical model that can help answer these questions.

II. A Theoretical Underpinning for the Debate

In order to define what constitutes discrimination and determine how such behavior might be detected, we must first understand how mortgage applications are underwritten. To do this, we use the mortgage underwriting model developed in Longhofer and Peters (1999).

Consider a world in which individuals want to buy a house, but lack sufficient funds to do so. As a result, they must obtain loans from a financial institution, which we will call a "bank." Individuals in this world are distinguished by two characteristics. The first is the likelihood with which they repay their loans, which we denote by $\theta \in [0,1]$. The second is their membership in one of two "groups," A or B. For ease of exposition, we will often refer to group A as the "white" group and to group B as the "minority" group. It should be clear, however, that these groups may alternatively be thought of as distinguishing individuals according to any publicly observable characteristic over which fair-lending laws might prohibit discrimination.

Let $g_A(\theta)$ represent the density of creditworthiness in the group A applicant pool and $g_B(\theta)$ represent the corresponding density for members of group B. It is important to note that these two densities generally will differ from one another.[4] We assume that both $g_A(\theta)$ and $g_B(\theta)$ are known to all banks.

Based on their costs of funds and the competitive interest rate in the market, banks determine a θ^* defining the minimum acceptable creditworthiness they are willing to approve. In a full-information world, lenders would know each individual's true θ and would approve only those applicants with $\theta \geq \theta^*$. Unfortunately, lenders cannot perfectly observe θ, but instead observe a signal s that is correlated with θ. This signal can be thought of as a summary of all the information a lender collects on a loan application, including the applicant's current income ratios and past credit history, the characteristics of the subject property, and so on.[5]

Let $p(s|\theta)$ denote the likelihood that a lender observes signal s from an applicant of type θ.[6] Using this signal generation process and the distribution of creditworthiness in the applicant pool, lenders derive an estimate of an applicant's expected creditworthiness, given his signal. We will often refer to this expected creditworthiness as an applicant's inferred "quality," which is denoted by

$$(1) \qquad q_i(s) = \int_{T_i} \theta \, \frac{p(s|\theta) g_i(\theta)}{\omega_i(s)} \, d\theta, \; i = A, B$$

where $\omega_i(s) = \int_{T_i} p(s|\theta) g_i(\theta) d\theta$, $i = A, B$, is the density of signals received from members of group i, and T_i is the set of group i applicants that apply for loans: $\{\theta \,|\, g_i(\theta) > 0\}$.

Expression (1) makes it clear that an applicant's inferred quality is simply his expected creditworthiness, where this expectation is taken over the lender's (Bayesian) posterior beliefs about the applicant's creditworthiness. Clearly,

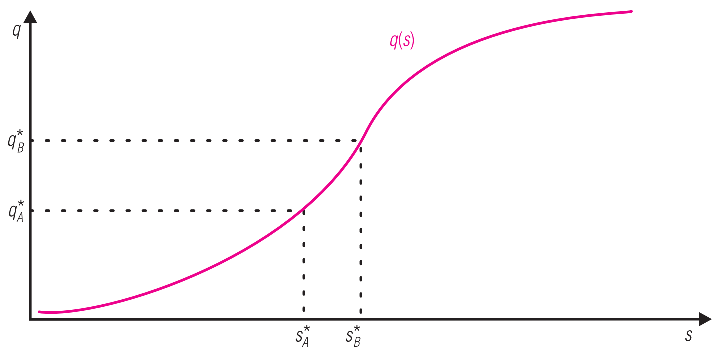
4 Longhofer and Peters (1999) analyze how individual self-selection can lead to endogenous differences in credit risk across groups. Alternatively, one could imagine that the distributions of credit risk in the underlying population differs across groups, leading to differences in these densities for the group applicant pools even in the absence of self-selection (see, for example, Calem and Stutzer [1995]).

5 Although an applicant's θ is assumed to be private information, the group to which he belongs can be costlessly observed by lenders, giving them an opportunity to discriminate on this basis if they so choose.

6 Although we assume that this signal generation process is the same for members of both groups, one could imagine a world in which it differed. Such a difference in $p(s|\theta)$ across groups lies at the heart of the cultural affinity hypothesis proposed by Calomiris et al. (1994).

FIGURE 1

Taste-Based Discrimination (Bigotry)



Bigoted lenders act on a taste for discrimination by holding minority applicants to a higher creditworthiness threshold ($q_A^* < q_B^*$). This implies that bigoted lenders will hold minorities to a more stringent underwriting standard than they do whites ($s_A^* < s_B^*$).

this posterior will be determined by both the distribution of credit risk in the overall applicant pool and the specific signal sent by the applicant. Given these beliefs, a lender will approve an applicant if and only if $q(s) \geq q^* \equiv \theta^*$.

Longhofer and Peters (1999) show that an applicant's inferred creditworthiness is increasing in his signal. That is, $q'(s) > 0$, confirming the intuition that applicants who send better signals tend to be more creditworthy. More importantly, this result assures us that there exists a unique cutoff signal, s^*, such that all applicants sending signals better than s^* are approved while those sending worse signals are not.

Defining Discrimination

The most striking omission in the debate over mortgage discrimination has been the lack of a formal definition of what behavior actually constitutes "discrimination." Perhaps this should not be surprising, given that the term is not defined in either the Equal Credit Opportunity Act or the Fair Housing Act—the two laws that directly prohibit discriminatory practices in the mortgage underwriting process. Presumably, this omission indicates Congress' belief that clarification would be superfluous because the meaning of the term is well understood by all reasonable people.

In our opinion, however, much of the controversy over how best to detect discrimination arises precisely because of differing implicit definitions of this central concept. To clarify the law's intent, we turn to *Merriam Webster's Collegiate Dictionary*, 10th ed., which defines discrimination as "the act, practice, or an instance of discriminating [making a distinction] categorically rather than individually." In other words, any difference in treatment across individuals based solely on group membership—rather than on personal characteristics specifically related to the performance of the loan—would constitute discriminatory behavior under the law. It is our interpretation (and the practice of the major bank regulators) that such a definition precludes lenders from applying group-specific underwriting standards, even if group membership is correlated with loan performance.[7]

Consistent with this idea, consider the following:

Definition: *A lender is said to* ***discriminate*** *against minority applicants if it requires them to meet a more stringent underwriting standard than it does white applicants; i.e., if* $s_B^* > s_A^*$.

When discrimination is explicitly defined in this way, it becomes clear that lenders' incentives to discriminate can arise for two distinct reasons. First, lenders may have Beckerian "tastes for discrimination," bigoted preferences that would manifest themselves through differences in the q^* required from members of each group. In other words, we say that a lender exhibits bigoted preferences if $q_B^* > q_A^*$. This source of discrimination is depicted in figure 1, which assumes that $q_A(s) = q_B(s)$ for every s.[8] In this case, the monotonicity of q implies that a bigoted lender will have an incentive to discriminate against members of group B by setting $s_B^* > s_A^*$.

Even without such tastes for discrimination, however, lenders may still have an incentive to discriminate if the overall pool of minority applicants is less creditworthy on average than the white applicant pool. In other words, if the minority applicant pool consists of relatively more low-θ individuals, $q_A(s)$ will lie above $q_B(s)$ for every s, giving banks an incentive for *statistical discrimination*.[9]

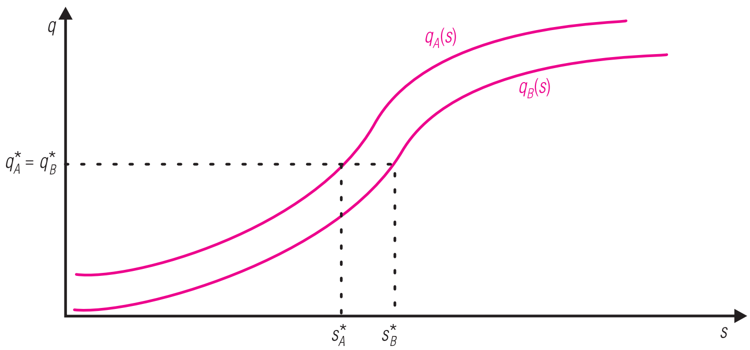
7 A noteworthy exception is the Equal Credit Opportunity Act, which specifically allows age to be considered in credit-scoring models, but only if it is to the advantage of older applicants.

8 This requires that the distribution of credit risk, $g(\theta)$, be the same across the two applicant pools.

9 Strictly speaking, it is only necessary that $q_A(s) > q_B(s)$ for $s < s_B^*$; Longhofer and Peters (1999) present a case in which statistical discrimination arises and yet $q_A(s) < q_B(s)$ for large s.

FIGURE 2

Statistical Discrimination



If minority applicants are less creditworthy on average than whites, this will be reflected in the lender's assessment of an applicant's inferred quality. Thus, $q_A(s)$ will lie above $q_B(s)$. In this case, even if lenders hold white and minority applicants to the same credit standard, $q_A^* = q_B^*$, they will have an incentive to statistically discriminate against minorities by setting $s_A^* < s_B^*$.

Statistical discrimination is depicted by figure 2, in which lenders set the same minimum creditworthiness standard for both white and minority applicants. Nevertheless, because minority applicants are less creditworthy (have lower θs) on average, lenders *believe* that a minority applicant sending any signal s is less creditworthy than a white applicant sending the same signal. Put another way, given the distribution of credit risk in their applicant pools, lenders know that a minority applicant sending any signal s is more likely to default than a white applicant sending the same signal. As a result, lenders have an incentive to hold minority applicants to a higher s^*.

Of course, these beliefs do not arise in a vacuum. As long as lenders accurately estimate the distribution of credit risk in their applicant pools, $g_i(\theta)$, the probability that a group i applicant repays his loan is $q_i(s)$. Thus, lenders that statistically discriminate find their beliefs validated ex post; on average, minority borrowers sending signal s_B^* repay their loans at exactly the same rate as white borrowers sending signal s_A^*, and both repay at the minimum rate acceptable to the bank.

Defining Default Rates

Given the above discussion, we can now define and analyze the different notions of default rates that have arisen in the mortgage discrimination controversy. As it turns out, participants in the debate rarely use the most useful of these notions, the *conditional default rate.*

Definition: *The* ***conditional default rate*** *for members of group i is the fraction of group i borrowers that actually default, conditional on their signal s:* $d_i(s) \equiv 1 - q_i(s),\ i = A, B.$

As discussed above, the likelihood that a group i borrower with signal s will end up defaulting on his loan is $1 - q_i(s)$. If there is a sufficiently large population of borrowers, then the actual fraction of group i borrowers sending signal s that default will be $d_i(s)$. From a lender's perspective, this conditional default rate is what matters most. After all, the lender's underwriting decision is ultimately determined by the likelihood that an applicant will default, given the information on his application. This is exactly what $d_i(s)$ measures.

Most of the mortgage discrimination debate, however, has focused on other notions of default rates.

Definition: *The* ***average default rate*** *is the fraction of all borrowers who actually default on their loans*:

$$(2)\qquad \bar{d}_i \equiv \int_{s \geq s_i^*} d_i(s) \frac{\omega_i(s)}{\int_{s \geq s_i^*} \omega_i(s)\, ds}\, ds,\ i = A, B.$$

Thus, the average default rate is simply the weighted average of the conditional default rates of individuals who were approved, where the weights are determined by the distribution of signals sent by approved applicants.

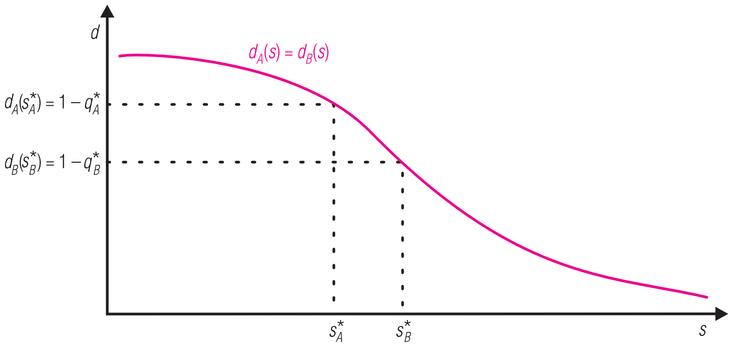
Finally, Berkovec et al. (1994), Calomiris et al. (1994), and Ferguson and Peters (1997, 1998) all consider varying notions of the "marginal" rate of default (or denial). Unfortunately, this concept has often been somewhat ill-defined.[10] It can be most precisely defined as follows:

Definition: *The* ***marginal default rate*** *is the fraction of defaulters among the borrowers sending the lowest approved signal:* $d_i^m \equiv d_i(s^*) = 1 - q_i(s^*),$ $i = A, B.$

10 A transgression to which both of the present authors plead guilty.

FIGURE 3

Default Rates under Bigotry



If credit risk is identically distributed across the minority and white applicant pools, bigoted lenders will require minority applicants to meet a more stringent underwriting standard (a higher s^*). Because the minority applicants who are rejected are those with the highest default risk, the average default rate will be lower for minorities than it will be for whites.

In the context of our model, the notion of a marginal default rate may be easily understood. In the real world, however, it is virtually impossible for regulators to identify the unique cutoff signal s^* used by lenders.[11] As a practical matter, then, the marginal default rate is often taken as either the average or the conditional default rate among borrowers sending signals below some arbitrary threshold. For example, Berkovec et al. (1994) analyze conditional default rates among FHA borrowers, under the presumption that they are riskier than those obtaining conventional loans; the authors further attempt to focus on "marginal" borrowers by dividing their sample into risk quartiles, comparing conditional default rates across white and minority borrowers in each quartile.

III. Deciphering the Debate

In section II, we introduced a simple theoretical model of the mortgage underwriting process and used it to help define discrimination and various notions of the default rate. In the present section, we use these definitions to reconsider the various arguments outlined in section I. In particular, we analyze the accuracy and the consistency of these arguments and consider why this debate has been so difficult to resolve. Finally, we discuss how our model's formal structure can be used to develop a better understanding of the underlying causes of any discrimination that does exist in the mortgage market. As we will argue, understanding the root causes of the problem is essential to designing effective policy responses.

The Default Rate Controversy

We begin by imagining a world where credit risk is identically distributed in both the minority and white applicant pools; that is, $g_A(\theta) = g_B(\theta), \forall\theta$. In such a world, the only possible source of discrimination is bigotry, which manifests itself as a higher required standard of creditworthiness for minorities ($q_B^* > q_A^*$). As shown in figure 1, lenders with such a taste for discrimination will respond by setting $s_B^* > s_A^*$. Nevertheless, the fact that $g_A(\theta) = g_B(\theta), \forall\theta$ implies that the inferred creditworthiness of a white applicant and a minority applicant, each of whom sends the same signal s, will be identical; that is $q_A(s) = q_B(s), \forall s$. Thus, conditional default rates will be the same across the two groups. But if conditional default rates are the same across groups, the average default rates of the two groups must diverge whenever banks set $s_B^* > s_A^*$; that is, $\bar{d}_A > \bar{d}_B$. Put another way, if minority and white borrowers are equally creditworthy on average, then any difference in the average default rate across groups must arise from lenders' tastes for discrimination.

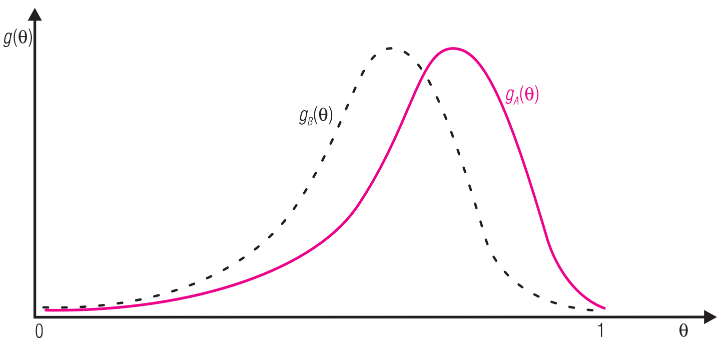
This idea is shown graphically in figure 3, in which conditional default rates are identical across the two groups. Because $q_B^* > q_A^*$, however, lenders set $s_B^* > s_A^*$, implying that white applicants with $s \in [s_A^*, s_B^*]$ are approved, while minority applicants sending the same signals are not. Ex post, these borrowers are the least creditworthy and hence the most likely to default. Because these applicants are excluded from the minority borrower pool, the average default rate of minorities who are actually approved for loans must be lower than that of whites.

Pursuing this logic, Becker (1993a), Brimelow and Spencer (1993), and Macey (1994) claim that Munnell et al.'s own data refute a conclusion of discrimination. After all, they argue, default rates across census tracts do not appear to be systematically related to the percentage of minorities residing in those tracts. While not conclusive, this evidence seems inconsistent with the notion that minorities default more frequently on average than whites.

11 For a discussion of how this model relates to econometric analyses of s^* by regulators and researchers, see Craig et al. (1998).

FIGURE 4

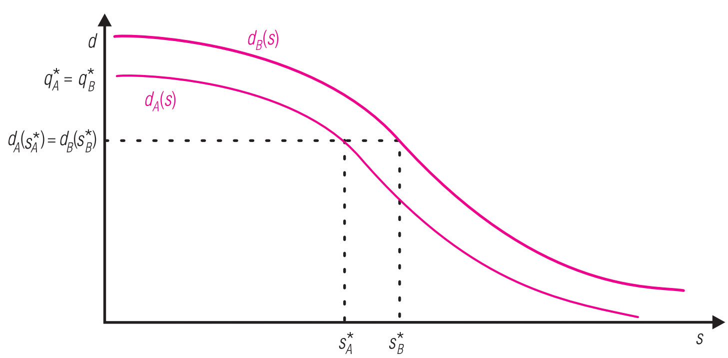
Differences in Underlying Creditworthiness across Groups



If minority applicants are less creditworthy on average than white applicants, the density of credit risk in the minority applicant pool will lie to the left of the density for the white applicant pool.

FIGURE 5

Statistical Discrimination and Default Rates



If minority applicants are less creditworthy than their white counterparts on average, lenders will have an incentive to statistically discriminate against them, even in the absence of bigotry. They do so by choosing s_A^* and s_B^* so as to equalize the inferred quality of the last applicant approved from each group $[q_A(s_A^*) = q_B^*(s_B^*)]$. This in turn implies that the marginal default rate of the two groups will be identical. However, the relative average default rates will be ambiguous.

Given their assumptions, Becker (1993a), Brimelow and Spencer (1993), and Macey (1994) are all correct in their conclusions. In our model, however, it is clear that for a group of borrowers, the average default rate depends not only on the underwriting standard applied to that group, but also on the distribution of signals sent by those borrowers, $\omega_i(s)$. This, in turn, depends on the distribution of credit risk in the group's applicant pool. Thus, ex ante differences in credit risk across applicant pools can lead to ex post differences in observed average default rates, even in the absence of discrimination.

This is the essential point made by Tootell (1993) and Galster (1993). They argue that Munnell et al.'s data verify that minority applicants are less creditworthy on average than their white counterparts.[12] This stylized fact is illustrated in figure 4, where $g_B(\theta)$ is shifted slightly to the left of $g_A(\theta)$.[13] In the context of our model, such differences in credit risk across groups cause lenders to adjust their assessment of an applicant's inferred quality, so that $q_A(s) > q_B(s)$, $\forall s$. As a result, if lenders were to hold both groups to the same s^*, the average default rate would be lower for white applicants. On the other hand, if lenders do set $s_A^* < s_B^*$, then the relative average default rate between the two groups will be ambiguous (see figure 5).

Thus we see that, if lenders are bigoted, there are two counteracting effects on minority default rates: the bigotry effect and lower average minority creditworthiness. On the one hand, lender tastes for discrimination imply that relatively more of the less-creditworthy white borrowers get included when calculating $\bar{d}_A$. On the other hand, for any given s, a minority borrower is more likely to default than a white borrower. Which of these two effects will dominate in practice is unclear.[14] As a result, Tootell and Galster both argue, average default rates cannot be used to disprove a charge of discrimination. In fact, if minority borrowers are less creditworthy on average than their white counterparts, a finding that both groups default at the same rate could only be consistent with lenders acting on bigotry against minorities.

Even though average default rates cannot disprove the existence of discrimination, other notions of the default rate may still be useful. The essential problem with average default rates is that they incorporate too much information. In contrast, conditional and marginal default rates focus on a specific subset of borrowers, making it easier to interpret what lies behind any differences across groups.

At its core, Becker's argument is driven by the fact that, if lenders hold minorities to a more stringent underwriting standard arising from bigotry, the "marginal" minority borrower

12 The Federal Reserve's National Surveys of Consumer Finances also support this conclusion.

13 Formally, one can think of the cumulative distribution function of credit risk in the white applicant pool, $G_A(\theta)$, as being first-order stochastic dominant over $G_B(\theta)$, the cumulative distribution of credit risk in the minority applicant pool; see Ferguson and Peters (1995).

14 In addition, the distribution of signals, $\omega(s)$, will differ across groups, further complicating this calculation.

TABLE 1

Relative Default Rates Implied by Bigotry and Statistical Discrimination

	Bigotry Only	Statistical Only	Both
Conditional default rates	$d_A(s) = d_B(s), \forall s$	$d_A(s) < d_B(s), \forall s$	$d_A(s) < d_B(s), \forall s$
Marginal default rates	$d_A(s_A^*) > d_B(s_B^*)$	$d_A(s_A^*) = d_B(s_B^*)$	$d_A(s_A^*) > d_B(s_B^*)$
Average default rates	$\bar{d}_A > \bar{d}_B$	$\bar{d}_A \geq\leq \bar{d}_B$	$\bar{d}_A \geq\leq \bar{d}_B$

(that is, the one with the lowest approved signal) will be more creditworthy than the marginal white borrower (see figure 3). In contrast, if lenders practice statistical discrimination, they do so precisely because they wish to ensure that the marginal white and marginal minority borrowers are equally creditworthy. In other words, statistical discrimination arises when lenders try to offset differences in the overall applicant pools across races (see figure 5). Thus, when we focus on marginal default rates, we see that Becker's argument remains valid, even in the presence of differential creditworthiness across racial groups.[15]

Unfortunately, regulators and econometricians cannot precisely observe a borrower's actual signal, nor can they pin down a lender's true underwriting guidelines. Thus, identifying the marginal borrower from each group is a practical impossibility. One way around this problem is to examine the default risk within a subset of borrowers whose creditworthiness is below some arbitrary threshold, as in Berkovec et al. (1994). Using a sample of FHA borrowers, they find that minorities in the highest-risk quartile have a significantly higher conditional default rate than do whites in the same risk class.[16] If we generically treat all borrowers in this risk class as "marginal," the results of Berkovec et al. would appear to suggest that statistical discrimination is the best explanation for those of Munnell et al.

Table 1 summarizes the default-rate implications discussed in this section. In a world in which members of both groups are equally creditworthy on average, conditional default rates will likewise be the same across groups [$d_A(s) = d_B(s), \forall s$]. In this case, if lenders exhibit tastes for discrimination the default rate of the marginal minority borrower will be lower than that of the marginal white borrower [$d_A(s_A^*) > d_B(s_B^*)$], causing the average default rate of minority borrowers to be lower than that of whites ($\bar{d}_A > \bar{d}_B$).

In contrast, when lenders have an incentive to statistically discriminate, minorities' conditional default rate will be higher than that of whites [$d_A(s) < d_B(s), \forall s$]; average default rates will be ambiguous. Nevertheless, lenders acting solely on a profit motive will discriminate against minorities only to the extent necessary to equalize the marginal default rate of members of each group [$d_A(s_A^*) = d_B(s_B^*)$].

Finally, when lenders act on both statistical and taste-based motives, conditional default rates will continue to be higher for minorities [$d_A(s) < d_B(s), \forall s$]. Interestingly, however, lender bigotry will cause the default rate of the marginal minority borrower to fall below that of the marginal white borrower [$d_A(s_A^*) > d_B(s_B^*)$], despite the higher conditional default rates for minority borrowers.

It is important to note that while conditional default rates may be useful in determining whether lenders have an incentive to discriminate, they cannot be used to determine whether lenders actually act on this incentive. Conditional default rates across groups will diverge whenever minority borrowers on average are substantially less creditworthy than white borrowers. This will be true whether or not lenders actually require minorities to meet a higher underwriting threshold, s^*. Thus, we must rely on denial rate analyses such as Munnell et al. to uncover discrimination and use default rate information to help determine the underlying source, if discrimination is determined to exist; we discuss the importance of this issue in more detail next.

What Constitutes Discrimination?

Taken together, Munnell et al.'s denial rate findings and Berkovec et al.'s default rate results pose a puzzle: How can marginal minority borrowers be significantly more likely to default than their white counterparts, yet still be less likely to be approved in the first place?

15 This point was made by Calomiris et al. (1994).

16 Although Berkovec et al. do control for numerous property and personal characteristics, an important limitation of their analysis is that their data do not contain any information on individual credit histories. The authors attempt to correct for any bias resulting from this omission, and still conclude that marginal minority borrowers default at a significantly higher rate than do whites.

The answer is statistical discrimination. In contrast to the taste-based discrimination considered by Becker, statistical discrimination is based on lenders' beliefs about applicant creditworthiness. Calomiris et al. (1994), Calem and Stutzer (1995) and Longhofer and Peters (1999) each show that statistical discrimination can occur if lenders' beliefs about creditworthiness differ across groups.[17] For example, Longhofer and Peters (1999) show that if a lender believes that minority applicants' average creditworthiness is lower than that of white applicants, then minority applicants will have to clear a higher hurdle —that is, minorities will be required to have better applications than whites—in order to receive a loan. Thus, although the inferred creditworthiness (the posterior assessment of creditworthiness) of white and minority borrowers is the same at the margin, econometric tests that use a borrower's application information (the signal) to proxy for his true creditworthiness will find that minorities default more often.

Although statistical discrimination provides a straightforward, plausible resolution of the results of Munnell et al. and Berkovec et al., many economists appear to be uncomfortable with this explanation. In fact, they seem to operate under the assumption that statistical discrimination is not, or should not be, illegal.

Where would such an assumption come from? Current law permits lenders to include most any underwriting variable that is demonstrably related to the profitability of a loan.[18] For example, lenders may consider an applicant's past credit history, even though this variable is highly correlated with race, because it is also strongly associated with an applicant's likelihood of repaying the loan. On the basis of this observation, it is reasonable to conclude that the proper scope of fair-lending law is to prohibit discrimination that arises for motives that are not profit-based, that is, from bigotry.

For example, Becker and Macey's focus on average default rates would seem rather misplaced, considering that average minority creditworthiness is lower than average white creditworthiness. Such a focus is perfectly reasonable, however, if one believes that statistical discrimination is not, in fact, discrimination at all. We are also struck by the lengths to which Munnell et al. go to reject the notion that their findings are driven by statistical discrimination: "The dearth of any evidence that minorities default more frequently, given their economic fundamentals, makes a conclusion of economically rational [statistical] discrimination problematic" (Munnell et al., [1996, p. 45]).

As we've already argued, however, existing fair-lending laws appear to prohibit both taste-based and statistical discrimination implicitly. Indeed, if statistical discrimination were permissible, lenders could simply include race as a variable in a statistically verified credit-scoring model; as long as the race variable showed a statistically significant impact on the loan's overall profitability, it would be permissible to consider race under such an interpretation of the law. Of course, this is expressly illegal. Furthermore, the "disparate impact test" in fair-lending enforcement effectively precludes the use of underwriting variables that themselves have no information content, but rather serve as proxies for race.

This discussion highlights why it is so important to confront issues like mortgage discrimination with a sound theoretical model. In the context of our formal model of the mortgage underwriting process, the definition of discrimination and the proper way of measuring it become clear. When both statistical discrimination and bigotry can result in differences in s^* across groups, the question of whether default rates can be used to detect discrimination becomes largely irrelevant.

This is not to suggest, however, that default rates are useless in the mortgage discrimination debate. In fact, one of the most important questions is how best to respond to any discrimination that does exist in the mortgage market. It is to this question that we turn next.

Uncovering the Source of Discrimination

Although borrower default rates are poorly suited for determining whether or not lenders discriminate against minority applicants, they can be useful in determining what incentives lie at the root of any discrimination that is uncovered. In other words, default rates may provide a means of distinguishing between statistical discrimination and bigotry.

It is worth asking why we are interested in identifying the source of mortgage market discrimination. After all, we have argued that this distinction is not important under the law. Nor is the root cause of discrimination likely to interest an individual victim. From a policy-

17 Arrow (1972a, 1972b) and Phelps (1972) offer the first discussion of statistical discrimination; they do so in the context of labor markets.

18 Such variables must pass the "disparate impact test," as we discuss below. See the Federal Financial Institution Examination Council's 1994 joint statement on fair-lending enforcement.

maker's perspective, however, it is quite important to know whether discrimination is belief-based or preference-based. Any policy's ability to eradicate discrimination will depend on the underlying source of that discrimination.[19]

For example, an appropriately designed penalty/subsidy scheme might counteract a bigoted lender's willingness to forgo profitable opportunities for lending to minority applicants, thereby inducing it to equalize its required q_A^* and q_B^*.[20] Alternatively (and perhaps more effectively), policymakers may attempt to combat taste-based discrimination by promoting competition in the mortgage market.

While these kinds of policies may work to eliminate bigotry, they can have the opposite effect on a statistical discriminator. Eradication of belief-based discrimination hinges on a policy's ability to make profit-maximizing lenders ignore costless information (the correlation between race and default) that, if used, would improve their profitability. Increased competition strengthens a lender's desire to employ such information, effectively increasing its incentives to statistically discriminate. Instead, statistical discrimination is best fought by focusing directly on the underlying source of differences across groups.

Likewise, penalty/subsidy schemes (much like the current use of the CRA as a tool for fair-lending enforcement) may be less desirable when the true source of discrimination is statistical correlation between race and creditworthiness. Although such a scheme may well enhance minorities' access to credit, it will do so at a very large cost: By making mortgage loans more attractive to less-creditworthy minority applicants, a penalty/subsidy solution would reduce the average creditworthiness of the applicant pool, *strengthening* the correlation between race and default. As a result, even larger subsidies or penalties would be required to eliminate statistical discrimination.

Finally, because statistical discrimination is motivated by a desire to maximize profits, in order to equalize s_A^* and s_B^* banks will have to hold whites to a higher creditworthiness standard than minorities ($q_A^* > q_B^*$). That is, to eradicate statistical discrimination, regulators must give banks a "taste for discrimination" against white applicants so that banks will have the incentive to pass up more profitable loans (to whites) in order to make less profitable ones (to minorities). Beyond its social and political implications, such a policy may be particularly difficult to implement.

Despite its importance from a policy perspective, uncovering the root cause of discrimination has not been the purpose of existing research. The reason is that both the statistical discriminator and the bigot will require a minority applicant to have a "better" application than a white applicant in order to be granted a loan. That is, a finding that $s_A^* < s_B^*$ is equally consistent with both preference-based and statistical discrimination.

Instead, the difference between taste- and preference-based discrimination is the fact that a bigot will require minorities to meet a higher minimum (inferred) standard of creditworthiness than whites (that is, $q_A^* < q_B^*$), while the statistical discriminator, because he has no taste for discrimination, is unwilling to forgo lending to creditworthy minority applicants and so holds minorities and whites to the same minimum credit standard (that is, $q_A^* = q_B^*$).

This fact suggests that conditional and marginal default rates may help determine the underlying source of any discrimination in the mortgage market. In particular, if conditional default rates are equal across groups, then statistical discrimination is an unlikely explanation for any discrimination that is uncovered through denial-rate analyses. On the other hand, if the conditional default rate of marginal minority borrowers exceeds that of marginal white borrowers, then bigotry becomes a less likely candidate for the root cause of discrimination.

Alternatively, Craig et al. (1998) propose a method for distinguishing empirically between statistical and preference-based discrimination that uses denial-rate data like those employed by Munnell et al. Unlike Munnell et al., they can address this issue because they take advantage of the structure inherent in the underwriting model developed by Longhofer and Peters (1999) and discussed here. Because the difference between statistical discrimination and bigotry lies in the relationship between q_A^* and q_B^*, the empiricist must be able to reconstruct the lender's Bayesian updating process (the underwriting process) to arrive at each applicant's inferred quality, $q(s)$. A finding that minority applicants are more likely to be rejected than whites, *conditional on* $q(s)$, would be consistent with preference-based discrimination. On the other hand, if the likelihood of being rejected varies across racial groups controlling for s, but does not vary after controlling for each applicant's q, statistical discrimination is the more plausible explanation.

19 See Longhofer (1995) for a related discussion.

20 Whether this involves the bank's raising its standard for whites or lowering it for minorities, however, is a more complicated question.

IV. Conclusion

We have shown here how the current mortgage discrimination debate has suffered because of inadequate theoretical underpinnings. Using a formal model of the underwriting process developed by Longhofer and Peters (1999), we are able to define what is meant by discrimination and to design tests for uncovering such behavior. Furthermore, we are able to define several different types of default rates precisely and to explore their implications under different underlying discriminatory incentives. Finally, we argue that developing appropriate policy responses to mortgage market discrimination depends crucially on understanding its root causes. Using our theoretical model, we are able to design some tests for uncovering this crucial information.

The definition of discrimination that arises out of our model implies that denial rate analyses like those performed by Munnell et al. are the proper tool for uncovering discrimination. This should not be interpreted, however, as suggesting that their results prove the existence of widespread discrimination in mortgage markets. Compelling as they are, we share some of the well-documented concerns about the veracity and interpretation of their results.[21] In the end, more research will be required to confirm or refute the existence of widespread discrimination and to understand its causes.

■ 21 See, for example, Horne (1994, 1997), Yezer et al. (1994), Hunter and Walker (1996), and Day and Leibowitz (1998). See also Browne and Tootell (1995).

References

Arrow, Kenneth J. "Models of Job Discrimination," in Anthony H. Pascal, ed., *Racial Discrimination in Economic Life.* Lexington, Mass.: Lexington Books, 1972a, pp. 83–102.

——**.** "Some Mathematical Models of Race Discrimination in the Labor Market," in Anthony H. Pascal, ed., *op. cit.*, 1972b, pp. 187–203.

Becker, Gary S. *The Economics of Discrimination,* 2d. ed. Chicago: University of Chicago Press, 1971.

——**.** "The Evidence against Banks Doesn't Prove Bias," *Business Week,* April 19, 1993a, p. 18.

——**.** "Nobel Lecture: The Economic Way of Looking at Behavior," *Journal of Political Economy,* vol. 101, no. 3 (June 1993b), pp. 385–409.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan. "Race, Redlining, and Residential Mortgage Loan Performance," *Journal of Real Estate, Finance, and Economics,* vol. 9, no. 3 (November 1994), pp. 263–94.

Brimelow, Peter, and Leslie Spencer. "The Hidden Clue," *Forbes,* January 4, 1993, p. 48.

Browne, Lynn Elaine, and Geoffrey M.B. Tootell. "Mortgage Lending in Boston—A Response to the Critics," Federal Reserve Bank of Boston, *New England Economic Review,* September/October 1995, pp. 53–78.

Calem, Paul, and Michael Stutzer. "The Simple Analytics of Observed Discrimination in Credit Markets," *Journal of Financial Intermediation,* vol. 4, no. 3 (July 1995), pp. 189–212.

Calomiris, Charles W., Charles M. Kahn, and Stanley D. Longhofer. "Housing-Finance Intervention and Private Incentives: Helping Minorities and the Poor," *Journal of Money, Credit, and Banking,* vol. 26, no. 3, pt. 2 (August 1994), pp. 634–74.

Craig, Ben R., Stanley D. Longhofer, and Stephen R. Peters. "Measuring Creditworthiness and Discrimination in Mortgage Lending," Federal Reserve Bank of Cleveland, unpublished manuscript, May 1998.

Day, Theodore E., and Stanley J. Leibowitz. “Mortgage Lending to Minorities: Where’s the Bias?” *Economic Inquiry*, vol. 36, no. 1 (January 1998), pp. 3–28.

Federal Financial Institutions Examination Council. “Policy Statement on Discrimination in Lending,” pt. III, *Federal Register*, vol. 59, no. 73 (April 15, 1994).

Ferguson, Michael F., and Stephen R. Peters. “What Constitutes Evidence of Discrimination in Lending?” *Journal of Finance*, vol. 50, no. 2 (June 1995), pp. 739–48.

——, and ——. “Cultural Affinity and Lending Discrimination: The Impact of Underwriting Errors and Credit Risk Distribution on Applicant Denial Rates,” *Journal of Financial Services Research*, vol. 11, nos. 1–2 (April 1997), pp. 153–67.

——, and ——. “Is Lending Discrimination Always Costly?” University of Illinois, unpublished manuscript, September 1998.

Galster, George C. “The Facts of Lending Discrimination Cannot Be Argued Away by Examining Default Rates,” *Housing Policy Debate*, vol. 4, no. 1 (1993), pp. 141–46.

Horne, David K. “Evaluating the Role of Race in Mortgage Lending,” *FDIC Banking Review*, vol. 7, no. 1 (Spring/Summer 1994), pp. 1–15.

—— . “Mortgage Lending, Race, and Model Specification.” *Journal of Financial Services Research*, vol. 11, nos. 1–2 (April 1997), pp. 43–68.

Hunter, William C., and Mary Beth Walker. “The Cultural Affinity Hypothesis and Mortgage Lending Decisions,” *Journal of Real Estate Finance and Economics*, vol. 13, no. 1 (July 1996), pp. 57–70.

Longhofer, Stanley D. “Rooting Out Discrimination in Home Mortgage Lending.” Federal Reserve Bank of Cleveland, *Economic Commentary*, November 1995.

——, and Stephen R. Peters. “Self-Selection and Discrimination in Credit Markets,” Federal Reserve Bank of Cleveland, Working Paper No. 9809, July 1998 (revised February 1999).

Macey, Jonathan R. “Banking by Quota,” *Wall Street Journal*, September 7, 1994, p. A14.

Munnell, Alicia H., Lynn E. Browne, James McEneaney, and Geoffrey M.B. Tootell. “Mortgage Lending in Boston: Interpreting HMDA Data,” Federal Reserve Bank of Boston, Working Paper No. 92–7, October 1992.

——, Geoffrey M.B. Tootell, Lynn E. Browne, and James McEneaney. “Mortgage Lending in Boston: Interpreting HMDA Data,” *American Economic Review*, vol. 86, no. 1, March 1996, pp. 25–53.

Phelps, Edmund S. “The Statistical Theory of Racism and Sexism,” *American Economic Review*, vol. 62, no. 4 (September 1972), pp. 659–61.

Tootell, Geoffrey M.B. “Defaults, Denials, and Discrimination in Mortgage Lending,” *New England Economic Review*, (September/October 1993), pp. 45–51.

Yezer, Anthony M.J., Robert F. Phillips, and Robert P. Trost. “Bias in Estimates of Discrimination and Default in Mortgage Lending: The Effects of Simultaneity and Self-Selection,” *Journal of Real Estate Finance and Economics*, vol. 9, no. 3 (November 1994), pp. 197–215.