Journal of Financial Services Research 11: 43–68 (1997)

Mortgage Lending, Race, and Model Specification

DAVID K. HORNE
Division of Research and Statistics Federal Deposit Insurance Corporation 550 17th Street, N.W. Washington D.C. 20429-9990

Abstract

This study examines the role of race in home mortgage lending by investigating the sensitivity of race estimates to variations in model specification. I compare parameter estimates based on a statistical model utilized by the Federal Reserve Bank of Boston, using a subset of the data that corresponds to FDIC-supervised institutions, with estimates obtained from several alternative variations specified to reflect information obtained from reviews of the mortgage loan application files. Estimates of the race effect are shown to be highly sensitive to the assumptions that underlie the model; minor modifications in model specification are sufficient to eliminate the race effect. The empirical results suggest that the statistical models used to evaluate the impact of race in mortgage lending may not provide reliable information about lending bias.

The perception that racial discrimination pervades the mortgage lending industry is attributable in large part to a widely cited econometric study of mortgage lending conducted by the Federal Reserve Bank of Boston. Munnell et al. (1992, 1996) examined mortgage lending in the Boston area and found that black and Hispanic applicants together were more likely to be denied loans relative to whites after accounting for all the factors considered by lenders. Several subsequent studies using the Boston mortgage application data (Carr and Megbolugbe 1993, Glennon and Stengel 1994, Hunter and Walker 1996) showed that the race effect persisted despite variations in model specification. Each of these studies claimed that the race effects uncovered by statistical analyses reflected lending bias. This paper focuses primarily on the Boston Fed study. However, the issues raised have relevance to the more general question of modeling the mortgage lending decision.

The Boston Fed study and its variants have been used to justify a wide range of government intervention in the mortgage market. The benefits of such programs depend on the extent of mortgage discrimination. These programs will be counterproductive if higher costs are associated with these programs because, in a competitive industry, such costs ultimately are passed on to consumers. Low- and moderate-income households, the intended beneficiaries of many of these programs, are most likely to be adversely affected by higher borrowing costs and the withdrawal of some lenders from the market. Lenders reporting high minority denial rates are frequently scrutinized by government agencies, the media, and community groups, but the focus on denial rates provides incentives for lenders to curtail community outreach programs that can generate higher minority rejection rates. The use of credit scoring models may become more popular as lenders attempt to eliminate the potential for subjective bias. However, the substitution of mechanistic models for lender judgment may exacerbate disparities in denial rates. Moreover, to the extent that studies create the perception of widespread lending discrimination, minority households may be deterred from considering home ownership as an option.[1] Thus, it is important that the studies that conclude that racial discrimination is prevalent in the mortgage industry be scrutinized carefully.

1. Models of mortgage lending discrimination

The Boston Fed conducted a comprehensive study of the role of race in mortgage lending, surveying mortgage lenders in the Boston area to supplement the 1990 Home Mortgage Disclosure Act (HMDA) data. Institutions in the Boston area reporting at least 25 mortgage applications were included in the study. The Boston Fed selected all black and Hispanic applicants reported by these institutions and a random sample of white applicants. The final sample consisted of 3062 applications for residential mortgage loans that were either approved or denied, of which 722 were applications from nonwhites and 2340 were applications from whites. Information was requested on 38 additional factors that mortgage underwriters and lenders indicated were important in evaluating applicant creditworthiness, including net worth, liquid assets, total assets, liabilities, proposed housing expenses and debt obligations, credit history, purchase price, loan amount, and appraised value. Data were collected on age, years of schooling, marital status, number of dependents, years employed in the line of work and in the current job for both applicants and coapplicants. The survey also addressed loan and property characteristics.

Nonwhite applicants in the Boston area experienced higher rejection rates. The Boston Fed reported that 28.1% of minority applicants were denied mortgage loans vs 10.3% of white applicants, excluding withdrawn applications and before controlling for other factors. The effect of minority racial status was estimated in a multivariate logit regression framework to control for other factors. Altogether, 13 variables in addition to an intercept term were included in the final logit model. The impact of race was found to be statistically significant, although the inclusion of other variables in the model reduced the size of the effect of minority status substantially. The study reported that ‘‘race does play a role as lenders consider whether to deny or approve mortgage loan applications. . . . A black or Hispanic applicant in the Boston area is roughly 60% more likely to be denied a mortgage loan than a similarly situated white applicant. This means that 17% of black or Hispanic applicants instead of 11% would be denied loans, even if they had the same obligation ratios, credit histories, loan-to-value, and property characteristics as white applicants’’ (Munnell et al. 1992, pp. 43–44). Based on these results, the authors concluded that lenders discriminated against minority applicants.[2]

In response to the Boston Fed study, the FDIC dispatched examiners to the 70 FDIC-supervised institutions that participated in the Boston Fed survey to evaluate lending patterns.[3] These institutions, which constitute more than half the 131 lenders surveyed by the Boston Fed, account for 45% of the mortgage applications in the Boston Fed sample. Examiners undertook an intensive review of individual loan files that corresponded to denied applicants (both white and nonwhite) who were more likely to have been approved than denied, according to the Boston Fed model, to search for evidence of racial discrimination. In addition, rejected applications by nonwhites were compared to approved applications by whites that appeared to be similar or weaker in terms of the reported factors in an attempt to identify possible instances of bias. A smaller number of applications by nonwhites and whites that were approved despite obvious financial weaknesses also were reviewed.

Analyzing evidence from the file reviews, Horne (1994) concluded that it was not possible to determine whether the race effect identified by the Boston Fed reflected differences in attributes considered by lenders but not captured by the statistical model or was caused by lending bias. Examiners were unable to verify the conclusions of lending discrimination reported by the Boston Fed. The file reviews revealed a number of important aspects of lending decisions that were not captured by the Boston Fed model. Four primary issues raised by the reviews were (1) the Boston Fed data contained numerous errors; (2) in many instances the dependent variable in question, application outcome, did not accurately reflect lender decisions or indicate the inability of an applicant to secure a mortgage; (3) a number of important factors influencing the ability to purchase a home were omitted or misspecified; and (4) the functional form of the model did not adequately reflect the underwriting process. Many of these problems appeared to understate the creditworthiness of white applicants relative to nonwhite applicants.

Horne (1994) did not quantify the impact of specification problems on the race estimates. However, a number of empirical studies analyzing the Boston Fed data concluded that the observed race effect reflected lending discrimination rather than specification problems because the parameter estimates were robust to variations in functional form. Munnell et al. (1992, 1996) demonstrated that the race effect was stable across several different models. Similarly, Carr and Megbolugbe (1993), Glennon and Stengel (1994), and Hunter and Walker (1996) showed that estimates of the race effect were insensitive to a number of other variations in functional form and data editing procedures.[4] The robustness of the race effect in many of the studies using the Boston Fed data is surprising, although it is possible to obtain stable parameter estimates that are seriously biased nonetheless. However, the stability of the parameter estimates is challenged by the empirical results presented in this paper.

Estimates of the race effect may be biased as a result of inadequate model specification (including omitted variable problems) or the use of simple, single-equation models to represent a system of complex, sequential decisions by applicants and lenders. Mortgage applicants do not represent a random sample of the population; the decision to apply is influenced by the amount of financial resources available as well as other factors, including the individual's self-assessed likelihood of obtaining mortgage financing. Applicants who expect their application would be denied may not apply (e.g., Maddala and Trost 1982). If individuals had perfect information about lending standards, all mortgage applications would be approved as a result of the self-selection process. Rosenblatt (1997) suggests that the variation in denial rates across groups may reflect differences in familiarity with lending standards. Information asymmetries may result from differences in experience with financial institutions, educational attainment, or exposure to relatives or peers who have successfully obtained a mortgage. Rachlis and Yezer (1993) note that application characteristics are not exogenous; marginal applicants may improve their likelihood of obtaining mortgage (adapting their original application or through subsequent revisions to their application) by increasing their down payment, adding a cosigner, or other actions.

Lenders may influence the distribution of loan outcomes by prescreening or otherwise discouraging particular classes of applicants (e.g., Benston 1981, Maddala and Trost 1982). Alternatively, lenders may encourage applications from particular groups through the use of community outreach programs, by participating in special lending programs, or by offering particular loan products. Rachlis and Yezer (1993), extending Maddala and Trost (1982), present a systematic treatment of the self-selection problem in a simultaneous equation framework and demonstrate how the problem of partial observability of this sequential process may introduce systematic bias into single-equation parameter estimates.

If data were available for each stage of the decision process for both applicants and lenders, one could estimate the structural form parameters to identify the role of race in the lending decision. Lacking such data, estimates of the race effect based on a single-equation model are subject to simultaneous-equations bias. Estimates of the extent of bias arising from the application of a single-equation model are produced by Yezer, Phillips, and Trost (1994) through the use of Monte Carlo simulations. They find that reduced-form models generate significant race effects even when the structural model is specified to be free of racial discrimination.

Although methodological problems with single-equation mortgage lending models have been discussed extensively in the literature (e.g., Barth, Cordes, and Yezer 1979, Benston 1981, Rachlis and Yezer 1993), the potential for systematic bias is not acknowledged by Munnell et al. (1992), Carr and Megbolugbe (1993), or Hunter and Walker (1996) and the problem is dismissed by Munnell et al. (1996). However, given the difficulties in estimating simultaneous models of sequential decisions with multiple unobserved factors, the use of single-equation models is not without utility. Due to the likely upward bias in race coefficients (such that the impact of minority status is exaggerated), Rachlis and Yezer (1993) argue that single equation models may be used to demonstrate the absence of lending bias, whereas if a race effect is obtained it may be impossible to determine whether the results reflect lending discrimination or other factors. Scrutiny of individual loan files may provide additional information about factors not captured by the model that are responsible for observed lending patterns.

2. Data

The analyses presented here focus on a subsample of the Boston Fed data, corresponding to FDIC-supervised mortgage lenders, for two reasons. First, the Boston Fed provided the FDIC with a complete data set only for this sample. Second, FDIC examiners could review loan files only at FDIC-supervised institutions. As a result, additional information on underwriting standards and on a portion of loan files at these institutions is available to me. In all, the FDIC-supervised institutions reported 1393 mortgage applications accounting for 45% of the entire Boston Fed data set.

The Boston Fed produced a list of applications that had been rejected for which the predicted denial probabilities, generated from their statistical model excluding race, were below 50%. This ''exception list'' was distributed to the federal regulatory agencies to assist in their investigations. Of the 70 FDIC-supervised institutions that participated in the Boston Fed study, 23 lenders reported no applications by nonwhites and another 24 lenders reported no rejections of nonwhite applicants. Of the remaining 23 institutions, four had no applications by nonwhites on the exception list. Examiners reviewed the 95 loan files corresponding to applications on the exception list at the remaining 19 institutions.

Day and Liebowitz (1994) documented many inconsistencies exhibited by the Boston Fed data. Comparison of mortgage payments with loan amounts yielded unreasonable imputed mortgage rates. Almost half the loans with loan-to-value ratios above 80% reportedly were not submitted for private mortgage insurance (PMI), although most of these were approved and PMI generally is required by lenders when loan-to-value exceeds 80%. Many applications with reported loan-to-value ratios above 95% were approved, although few lenders originate such loans. Monthly income on an annualized basis was substantially different from reported annual income in many instances. The reported housing expense-to-income ratio differed greatly from the housing expense ratio calculated from housing expense and income data. Calculated net worth was reported to be substantially negative for a large number of applicants, many of whom received mortgage loans. These and other data problems also plagued the sample of applications reported by FDIC-supervised institutions.

Inspection of applicant files at FDIC-supervised lenders confirmed the existence of a large number of data errors and revealed several systematic reporting errors (Horne 1994). Some lenders reported information directly from the original applications without incorporating information subsequently collected during the processing of the application. The application data often were inaccurate as lenders determined that financial information was exaggerated or that important information could not be verified. In addition, lenders appeared to be confused by survey instructions, which required lenders to alternate reporting data in dollars and in thousands of dollars. Survey questions 35 and 36 requested information on liquid assets and total assets in thousands of dollars. The next question asked lenders to report total monthly nonhousing debt obligations in dollars. The following question (38) requested data on total liabilities in thousands of dollars. Given the large volume of information requested from lenders, coding errors may have resulted as lenders alternated between reporting in dollars and thousands of dollars. It is impossible to determine precisely how many of the incongruous observations occurred as a result of reporting errors. Examples of errors that appeared to involve scaling problems include four instances in which calculations using proposed housing expense and income produced reported ratios that were off by a factor of exactly 100.[5] Similarly, liabilities appeared to be reported in dollars rather than in thousands of dollars.[6] As a result, 69 applications were observed to have negative net worth (calculated from reported assets and liabilities); nevertheless, 57 of these applications were approved. In some cases the negative net worth amounted to hundreds of thousands or even millions of dollars.

Where possible, I corrected the data based on reported information as well as material obtained from the file reviews. The data corrections made to the sample are summarized in table 1. The file review data in conjunction with the survey data revealed that 71 applications were improperly included. Twenty-one were applications for refinancing, home improvement, or construction loans or had actually been withdrawn, all of which

Table 1. Modifications to Boston Fed sample

Sample Modifications	Sample Size	% of FDIC Sample
Original Boston Fed model[a]	3062	
Sample of FDIC-supervised lenders[b]	1393	100.0
Delete construction, withdrawn, refinance, investor applications	66	
Delete overqualified applications	5	
Delete loan-to-value < 30%	40	
Revised sample 1[c]	1282	92.0
Delete ambiguous outcomes	61	
Revised sample 2[d]	1221	87.0

[a]Corresponds to Table 4, model 1.
[b]Corresponds to Table 4, model 2.
[c]Corresponds to Table 4, model 3, all models in Table 5, Table 6, models 1 and 2.
[d]Corresponds to Table 6, models 3 and 4.
Note: Changes to the revised sample 1 include modification of housing expense and obligation ratios (38 observations); to loan amount, purchase price, and appraised value (16 observations), and to loan outcome (4 observations corrected).

were supposed to have been excluded from the Boston Fed analysis, which was limited to applications for conventional home mortgages. Another 45 applications were from investors (properties not owner occupied) and hence also should have been excluded. Five additional observations were deleted because they represented applicants who were overqualified for participation in special mortgage loan programs targeted to low-income households. In such circumstances, a denial does not indicate that a lender is unwilling to provide financing under a conventional mortgage program.

Examination of the data revealed serious inconsistencies for many observations with low loan-to-value ratios. The loan amounts were unreasonably small in many cases, proposed housing expenses were inconsistent with the loan amounts, and other problems suggested that the applications with very low loan-to-value ratios contained unreliable data. Moreover, because lenders have little risk in making loans with such low loan-to-value ratios, the underwriting for these types of loans should differ materially from that applied to more typical mortgage applications. Consequently, 40 applications with reported loan-to-value ratios below 30% were deleted from the sample.[7]

Modifications were made to correct for bad data where the errors appeared to result from reporting confusion with respect to the units of measure, including housing expense and obligation ratios (38 observations), as well as for loan amount, purchase price, and appraised value (16 observations); much of this information was provided by examiners as a result of the file reviews. Four additional observations incorrectly reported as denials were recoded as approvals. A total of 58 observations were modified. Altogether, 1282 observations remained in the revised data set (revised sample 1).

In some instances, denials did not necessarily indicate that lenders were unwilling to provide mortgage financing. Overqualified applicants were excluded from the sample, as noted previously. The proper treatment of other denial categories is less clear. Counteroffers made by lenders but rejected by applicants are reported as lender denials under HMDA. Counteroffers accepted by applicants, in contrast, are reported as lender approvals. Although lenders could use counteroffers to discourage minority applicants, counteroffers also may be a means for lenders to accommodate applicants who have potentially acceptable applications or who apply for mortgage products not offered by them. Applications submitted to, and rejected by, private mortgage insurance companies also are reported as lender denials. The file reviews uncovered other instances where lenders appeared willing to provide mortgage financing but the outcomes were reported as denials because the transactions were precluded by title problems or housing code violations.[8] The 61 observations with questionable outcomes were first treated as denials and subsequently were recoded as approvals in the revised sample 1. Finally, these 61 observations were deleted to produce the revised sample 2, which consisted of 1221 observations.

The revisions to the sample had a relatively small impact on the sample attributes, as the data in table 2 demonstrate. The columns corresponding to each racial group show the same information for the sample of FDIC-supervised lenders before and after the data editing just described (for revised sample 1). Applications by whites differed from

Table 2. Applicant attributes, by race, as a proportion of each demographic group, (original sample of 1393 observations vs. revised sample of 1282 observations)

Characteristics	Nonwhite (original)	Nonwhite (revised)	White (revised)	White (original)
Denials	33.8%	30.1%	9.2%	9.6%
Credit variables				
Insufficient credit history	9.2%	8.8%	2.2%	2.0%
Delinquent credit history	11.6%	11.8%	6.1%	6.0%
Serious delinquencies	13.0%	12.9%	6.9%	6.8%
Public record of defaults	14.0%	14.7%	4.7%	4.7%
Meets credit standards	74.1%	75.0%	93.5%	93.3%
Unverified information	16.0%	14.3%	3.8%	4.2%
No mortgage payment history	81.9%	82.4%	66.2%	64.3%
1 or 2 late mortgage payments[a]	1.4%	1.5%	1.8%	1.8%
> 2 late mortgage payments[a]	0.7%	0.7%	1.2%	1.5%
Applicant characteristics				
Total monthly income (median)	\$2747	\$2734	\$3550	\$3565
Net worth (median, \$000)	\$38.0	\$37.4	\$82.7	\$88.0
Liquid assets (median, \$000)	\$17.5	\$17.0	\$35.0	\$35.9
Liquid assets $\geq$ closing costs	55%	55%	65%	65%
Years of school (applicant) (mean)	14.3	14.2	15.5	15.4
Application characteristics				
Down payment (median)[b]	15%	14%	22%	23%
Housing expense/income (median)	27%	27%	26%	26%
Total obligations/income (median)	34.8%	35.0%	33.0%	33.0%
Denied PMI	9.6%	10.3%	2.6%	2.2%
Percent of 2–4-family homes	24.9%	26.1%	8.9%	9.5%
Number of applicants	293	272	1010	1100

[a]Percent of homeowners.

[b]Appraised value minus loan amount (as a proportion of appraised value).

applications by nonwhites in a number of important respects.[9] A larger proportion of nonwhite applicants had insufficient credit history. Nonwhite applicants were more likely to have a history of credit delinquency and a public record of defaults. Nonwhite applicants were less likely to meet a lender's credit standards, and lenders reported more problems verifying information reported by nonwhite applicants. Based on the mortgage payment history variable, a larger proportion of nonwhite applicants did not already own a home, as indicated by having no mortgage payment history (82.4% vs. 66.2% of whites), and nonwhite homeowners were more likely to have a record of late mortgage payments.

On average, white applicants reported higher levels of income, net worth and liquid assets. The median income of whites exceeded median income of nonwhites by 30%. Nonwhite applicants reported a median net worth (calculated as total assets minus liabilities) of $37,400 vs. $82,700 for whites. Liquid assets were higher for white applicants; median liquid assets for whites exceeded that of nonwhite applicants by a factor of 2.[10] Liquid assets alone were sufficient to meet projected closing costs (as discussed in more detail later) for 65% of white compared to 55% of nonwhite applicants. White applicants completed an average of 1.3 additional years of school relative to nonwhite applicants.

Down payments were higher for white applicants, but there was little difference in the average housing expense and total obligation ratios (each measured as a proportion of income) between white and nonwhite applicants. PMI was denied for nonwhite applicants almost four times as often as for whites. Differences in the types of properties represented by the mortgage applications also were evident. Nonwhite applicants were more likely to apply to finance 2–4 family properties than white applicants.

3. Model specification

The impact of race in mortgage lending is generally estimated in a multivariate regression framework to control for the factors considered by mortgage underwriters when evaluating mortgage applications. Because the dependent variable, loan application outcome, is a dichotomous variable, qualitative response models are used to estimate the impact of the exogenous variables ($\mathbf{X}$) on the probability that the application will be denied. The probability that application i will be denied, which can be expressed as $P(Y_i = 1)$, is a function of the applicant's attributes and the parameter estimates (β) generated by the logit model: $P(Y_i = 1) = \exp(\beta' \mathbf{X}_i)/(1 + \exp(\beta' \mathbf{X}_i)$.

Reported HMDA data indicate that denial rates for nonwhite applicants are higher than for white applicants. Statistical models typically account for some, but not all, of this disparity by controling for the other factors. However, the estimated race effect will be biased if important factors that underwriters consider are omitted from (or inadequately captured by) the model and these factors are correlated with race. As the data in table 2 demonstrate, many of the factors that relate to creditworthiness are correlated with race.

Many problems with mortgage lending models have been discussed in the literature. Three particular specification issues that have received limited attention but which are relevant to the analysis are considered in the following specification. These include interpretation of the outcome variable, measurement error, and functional form.

3.1. Model specification issues

The assertion that higher denial rates for nonwhites indicate discrimination by mortgage lenders presumes that the loan outcomes reflect lenders' willingness to provide mortgage financing. This assumption may be inappropriate for a number of reasons. In some instances, lenders may not provide financing because private mortgage insurance could not be obtained. This issue is particularly important for lenders who originate mortgages for resale in the secondary market, where mortgage insurance is required. If the submission for mortgage insurance implies that the lender is willing to provide financing (subject to mortgage insurance approval), it would be inappropriate to interpret applications rejected by mortgage insurance companies as lender denials. Similarly, applicants who apply for special affordable housing programs may be denied by lenders in accordance with program regulations if their income or assets exceed program limits. Such applicants may be eligible for conventional mortgage loans. In sum, it is inappropriate to attribute higher minority denial rates to discrimination by mortgage lenders if application outcomes do not reflect lenders' decisions.

Measurement error is a serious problem in mortgage lending models. Income perhaps is the most visible underwriting factor and is reflected in the housing expense and total obligation ratios (proposed housing expenses and total obligations are evaluated as a proportion of income). Although statistical models generally include current income, lenders in fact are concerned with applicants' ability to meet mortgage payments in the future. In this context, current income may be considered an indicator of permanent income, which is unobserved. To obtain a better gauge of permanent income, lenders consider employment stability, job tenure, promotion record, educational attainment, and other measures generally referred to as *compensating factors*. Such factors are difficult to incorporate into statistical models, but omitting these factors may bias the model results.

Lenders examine applications to determine whether applicants have sufficient funds to finance the down payment and other closing costs. Applicants may rely on liquid assets to finance these costs. However, applicants also may obtain funds from other sources. Applicants who currently own property often rely on the proceeds from the sale of the property to meet financing requirements. First-time home buyers often depend on gifts from relatives and other sources of funds such as borrowing from retirement accounts. The Boston Fed did not collect data indicating whether applicants had sufficient funds to meet closing costs. No data on the amount of gifts provided by relatives were collected; home equity was not measured directly (but would be included in the net worth measure). Net worth excludes gifts but includes items (such as personal property) that are not particularly relevant to lending decisions. As a result, it is impossible to determine from the data whether applicants had sufficient funds to qualify for a mortgage.

Distinctions that appeared to be important to underwriters were not captured by the Boston Fed's credit risk variable. If an applicant's credit history is less than satisfactory, lenders may consider the circumstances associated with past credit problems to gauge creditworthiness. If a credit report reveals evidence of past problems that since have been resolved, reasonable explanations may be accepted by lenders. Such one-time events as medical difficulties, educational expenses, or family problems might not be considered obstacles to obtaining a mortgage if the debts subsequently were settled. Indeed, the resolution of such credit problems may be interpreted as indicating a willingness to meet financial responsibilities. The credit risk variable does not distinguish between one-time credit problems resolved in the past and persistent credit problems that involve current accounts, although lenders appear to treat these cases differently. Lenders did report whether applicants met the institution's credit standards, a variable that may reflect distinctions important to lenders but not captured by the other credit risk variables. However, this variable was excluded from the Boston Fed model.

The regression models used to analyze lending decisions measure the relative importance of each factor independent of the values of other factors. In fact, underwriting is a sequential process. Lenders first examine a small number of factors, such as the loan-to-value ratio, proposed housing expense and obligation ratios, and credit history. Although the model specification implies that the probability of denial is a function of all the factors in the model, a major problem with just one of the primary underwriting factors (such as the housing expense or debt ratios or credit reports) may generate a denial. In this case the values of the other variables become irrelevant. Indeed, if denial is warranted based on examination of preliminary data, not only will some information not be verified, but other data, such as the appraised value of the property, may not be collected, further complicating model estimation. Examiners noted such problems in the review of the application files. For marginal applicants, the lender normally will assess additional data (the compensating factors) to tip the decision one way or the other. The relative importance placed on these compensating factors is variable and depends on the values of other variables in a very nonlinear manner. If creditworthiness is measured with error (because some files are incomplete or important compensating variables are omitted) and measurement error is correlated with the application outcome, the estimated race effect can be systematically biased.[11]

Although problems with interpretation of the dependent variable, measurement error and functional form complicate model estimation, the race effect might not be systematically biased if these errors were random. However, this assumption is unwarranted. The potential for systematic bias may be illustrated with a simple example for each of these problems. In the Boston Fed sample approximately 30% of nonwhite applicants apply for mortgage financing through affordable housing programs vs. just under 7% of white applicants. The potential for denial of overqualified applicants therefore is greater for nonwhite applicants, introducing a serious problem for estimation because the model implies that lenders are likely to deny better-qualified nonwhite applicants. The ability of white applicants to finance the down payment and closing costs may be underestimated to the extent that whites have greater access to financial assistance from relatives and are more likely to rely on equity from the sale of property.[12] Consequently, white applicants tend to have funds that permit them to make larger down payments. Lenders may attach less importance to other factors given relatively high levels of owner equity because owners have greater incentive and ability to meet their mortgage obligations (which are lower when down payments are higher, *ceturis paribus*), and lenders have more collateral value as protection against foreclosure. To the extent that such factors correlated with race are not captured by other variables, their influence will show up in the estimated race effect.

3.2. The Boston Fed model specification

The variables included in the models are described in table 3. Most variables are defined to be consistent with those in the Boston Fed study. The dependent variable (OUTCOME) is a dichotomous variable indicating whether the application is approved or denied. RACE is defined to indicate whether the applicant was black or Hispanic, or white. The front-end ratio (proposed housing expense/income) is captured by a dummy variable (HOUSE) that indicates whether the ratio exceeds 30%. The back-end ratio (DEBT) is captured by a continuous variable calculated as total debt and proposed housing payments as a% of income. Net worth (NTWORTH) measures total net worth. A dummy variable (LQASSET) is specified to indicate whether the applicant had sufficient liquid assets to meet closing costs. This variable is a proxy for whether applicants had the necessary amount of funds available to pay closing costs, estimated as the down payment plus two months mortgage payments, but does not reflect potential sources of funds other than liquid assets. Some of the models presented in the Boston Fed study utilized liquid assets, but there was no variable indicating whether applicants had sufficient funds to meet

Table 3. Definition of variables

Variable Name	Variable Definition
OUTCOME[a]	1 if denied; 0 if approved
RACE[a]	1 for black or Hispanic applicants; 0 for whites
HOUSE[a]	1 if housing expense/income > 0.30; 0 otherwise
DEBT[a]	Total debt payments/income (%)
NTWORTH[a]	net worth ($000)
LQASSET	1 if liquid assets $\geq$ closing costs; 0 otherwise
LTV[a]	Ratio of loan amount to appraised value
LTV91	1 if 91% $\leq$ LTV $<$ 96%; 0 otherwise
LTV96	1 if 96% $\leq$ LTV $<$ 100%; 0 otherwise
LTV100	1 if LTV $\geq$ 100%; 0 otherwise
PUBREC[a]	1 if public record of defaults; 0 otherwise
NOMORT	1 if no mortgage payment history; 0 otherwise
LATEMORT	1 if more than two late mortgage payments; 0 otherwise
MORTPAY[a]	Mortgage payments (continuous variable ranging from 1 = no late pay to 4 = more than two late payments)
NOCREDIT	1 if insufficient consumer credit history; 0 otherwise
DELCREDIT	1 if delinquent credit (60+ days); 0 otherwise
BADCREDIT	1 if serious credit problems (90+ days); 0 otherwise
CREDIT[a]	Consumer credit (continuous variable ranging from 1 = no slow pay to 6 = serious delinquencies[a])
PMI[a]	1 if denied PMI; 0 otherwise
RENT[a]	Rent/property value in census tract
2–4 FAMILY[a]	1 if 2–4-family housing; 0 for single-family housing
SELFEMP[a]	1 if self-employed; 0 otherwise
UNEMP[a]	Industry-specific unemployment rate
EDUCAT	Applicant's years of education
UNVERIF	1 if application data could not be verified; 0 otherwise
STANDARDS	1 if applicant did not meet lender's credit standards; 0 otherwise

[a]Variables included in Boston Fed model.

closing costs. The loan-to-appraised-value ratio is captured as a continuous variable (LTV) and as a series of dummy variables for different LTV ranges as defined in table 3. If collateral problems were noted in the file reviews or the appraised value was unavailable, the property value was set equal to the loan amount.

Several variables were created from mortgage payment and consumer credit history information collected by the Boston Fed. A variable indicated whether there was a public record of defaults, charge-offs, collection actions, or bankruptcies (PUBREC). Dummy variables were created to indicate whether there was no mortgage history (NOMORT) or more than two late mortgage payments (LATEMORT). A mortgage variable duplicating that used by the Boston Fed was also created. This variable (MORTPAY) ranged from 1 (for no late payments) to 4 (more than two late payments). Similar variables were created to represent consumer credit. An insufficient consumer credit history was reflected by the variable NOCREDIT. Other variables indicated whether accounts had been identified as 60–89 days past due (DELCREDIT) or at least 90 days past due (BADCREDIT). In addition, a consumer credit variable (CREDIT) was created with values ranging from 1 (no slow pay accounts) to 6 (serious delinquencies) to reproduce the credit variable used by the Boston Fed.[13]

Dummy variables were created to indicate whether private mortgage insurance was denied (PMI). The RENT variable reflected the average rental value as a proportion of property values for each census tract. Additional variables indicated whether the property represented a multifamily unit (2–4FAMILY). SELFEMP reflected whether the applicant was self-employed. UNEMP, also created by the Boston Fed, was a measure of the unemployment rate associated with the industry in which the applicant was employed. Three additional variables collected by the Boston Fed but not included in the models were incorporated into the models. EDUCAT was defined as years of education. UNVERIF indicated whether the lender could verify application data. The STANDARDS variable was based on the lender's assessment of whether the applicant met the lender's credit standards.

The models presented in table 4 all have the same functional form to facilitate comparison of results based on different data sets. The first column in table 4 reproduces the parameter estimates reported in Munnell et al. (1992). The second set of parameter estimates in table 4 show the results of the same model applied to the 1393 applications from FDIC-supervised lenders prior to any editing. The third column shows the parameter estimates generated when the same functional form is applied to the revised sample 1 consisting of 1282 observations (see table 1). The primary modifications to the data include the exclusion of withdrawn applications, construction loans, refinancing, and investment properties, recoding observations incorrectly reported as denials (including overqualified applicants), and deletion of observations with loan-to-value ratios below 30%.

The coefficients generated by applying the Boston Fed model to the two FDIC samples generally are similar to those reported by Munnell et al. (1992, 1996). The potential for discrimination is indicated by the statistical significance of the race coefficient. The coefficients in a logit model have no intuitive interpretation, but the influence of race on the denial probabilities can be calculated.

Estimating the impact of race is complicated by two factors. First, groups comparisons are usually based on differences in denial rates. However, logit models generate denial

Table 4. Determinants of mortgage loan denials (Boston Fed model specification)

Variables	Boston Fed Model (original coefficients)	Boston Fed Model (preedited FDIC sample)	Boston Fed Model (revised sample 1)
RACE	0.68	1.12	0.67
	(0.0001)	(0.0001)	(0.0058)
INTERCEPT	−6.61	−6.74	−8.63
	(0.0001)	(0.0001)	(0.0001)
HOUSE	0.47	0.53	0.43
	(0.0014)	(0.0220)	(0.0938)
DEBT	0.04	0.04	0.05
	(0.0001)	(0.0001)	(0.0001)
NTWORTH	0.00008	0.00009	0.00014
	(0.2713)	(0.1834)	(0.0800)
LTV	0.58	0.51	2.39
	(0.0014)	(0.0055)	(0.0008)
PUBREC	1.2	1.3	1.57
	(0.0001)	(0.0001)	(0.0001)
MORTPAY	0.35	0.62	0.60
	(0.0027)	(0.0003)	(0.0028)
CREDIT	0.33	0.32	0.32
	(0.0001)	(0.0001)	(0.0001)
PMI	4.70	4.72	4.48
	(0.0001)	(0.0001)	(0.0001)
RENT	0.68	−0.02	0.03
	(0.0005)	(0.9747)	(0.9719)
2–4-FAMILY	0.58	0.58	0.65
	(0.0003)	(0.0221)	(0.0252)
SELF-EMP	0.52	0.15	0.06
	(0.0051)	(0.5956)	(0.8661)
UNEMP	0.09	0.06	0.04
	(0.0009)	(0.1540)	(0.3937)
AIC	—	770.2	635.0
Somers's *D*	—	0.75	0.79
Sample size	3062	1393	1282
Missing observations[a]		0	0

Note: p-values are reported in parentheses; the p-values in column 1 were calculated from the t-statistics reported by Munnell et al. (1992).

AIC (Akaike information criterion): $-2 \log L + 2(k + s)$, where L is the likelihood function, k is the number of ordered values for the dependent variable (loan outcome in this case), and s is the number of explanatory variables in the regression.

Somers' D statistics: The number of concordant pairs minus the number of concordant pairs as a proportion of all pairs with different outcome values.

[a]Observations will be excluded from the regression if data for any variable in the model is missing.

probabilities for each applicant (i.e., $P(Y_i = 1)$) rather than outcomes, so denial frequencies cannot be calculated directly. The mean of the individual probabilities for each group (where each individual's estimated denial probability is a function of their attributes and the logit parameter estimates) are constrained in logit models to equal the actual denial rates for each group (Maddala, 1983). Thus, the effect of changes in racial status on denial rates can be estimated for each racial group by examining the changes in the mean denial probabilities associated with different values of the race variable. Second, the estimated denial probabilities are a nonlinear function of the estimated coefficients and the values of the exogenous variables. As a result, the effect of a change in racial status will vary for each group because the attributes vary with race. Because of these nonlinearities, the effect of racial status (i.e., the change in denial rates associated with different values of the race variable) is calculated separately for nonwhite and white applicants. In effect, the actual minority denial rate is compared to the denial rate that would be expected had these same applicants been treated as whites, and vice versa.

The race coefficient of 1.12 (from the second model in table 4) implies that 21.2% of applications from nonwhites in the sample would have been denied had they been treated as applications from whites, compared to an actual (raw) minority denial rate of 33.8% for this sample.[14] Similarly, 19.1% of applications from whites would have been denied if they had applied as nonwhite applicants (holding other attributes constant), compared to the actual (raw) 9.6% denial rate. Thus, the disadvantage associated with minority status implied by the parameter estimates ranges from 12.6 percentage points (for minorities) to 9.5 percentage points (for whites). The race coefficient of 0.67 in the last column of table 4 (the Boston Fed model utilizing the revised sample 1), in turn, implies a disadvantage associated with minority status from 6.7 percentage points (for nonwhites) to 4.6 (for whites). These numbers are consistent with the results reported by Munnell et al. (1992, 1996), who identify a race effect of 6 percentage points in their working paper (table 5) and 8.2 percentage points in their published paper (table 2).

The racial disparity in denial rates for the sample of FDIC-supervised lenders exceeds that for other lenders participating in the Boston Fed study. The minority denial rate in the sample of FDIC-supervised lenders of 33.8% (see table 2) is greater than the 28.1% denial rate reported by Munnell et al. for all lenders, and the denial rate for whites was lower in the FDIC sample (9.2% compared to 10.3% for the complete data set), although the FDIC sample is a subset of all lenders included in the Boston Fed study. Correcting for errors reduced the race effect estimated within the framework of the Boston Fed model considerably. If similar errors affect the data for non-FDIC lenders and were corrected by the Boston Fed, the race effects reported by Munnell et al. would be lower.

Munnell et al. (1992) focus on relative differences in denial rates, reporting that minority status increases the probability of denial by about 60% and that denial rates for nonwhites range from 1.4 to 2.7 times that for whites, based on parameter estimates derived from single-race equations. The emphasis on relative denial rates obfuscates comparisons between groups because the scale is sensitive to the level of denial rates. Most observers would agree that lending data would generate little attention if 99% of whites were approved, vs. 98% of nonwhite applicants. However, the denial rate for nonwhites in this case would exceed that for whites by 100% (2 vs. 1%). For this reason, the discussion in this paper focuses on the absolute differences in denial rates.

Changes in the functional form of the model and variations in the sample attributes influence model fit. A model that more closely represents the lending decision is likely to fit the data better. To compare the results of different models, two goodness-of-fit measures are provided. The Akaike information criterion (AIC) offers a simple statistic for comparing competing models. This statistic incorporates an adjustment for differences in degrees of freedom, particularly useful as the models presented here generally are not nested and include different numbers of variables. The model with the smallest AIC value is preferred, as a lower value reflects a better model fit.[15] The second measure reflects the rank correlation between the observed loan application outcome values and the predicted probabilities. Each pair of responses with different loan application outcomes is compared to produce this statistic. A pair of observations with different outcome values is concordant (discordant) if the larger response $(Y_i = 1)$ has a higher (lower) predicted value than that observation associated with the smaller response $(Y_j = 0)$. The reported measure of predictive fit, the Somers's *D* statistic, is defined as the number of concordant pairs minus the number of discordant pairs, divided by the total number of pairs with different responses. Comparisons of the AIC and Somers' *D* statistics show that the model fit and the predictive ability of the estimates using the revised data are superior to the previous model using unedited data. The data corrections improve the model fit and reduce the estimated race effect.

3.3. Alternative model specifications

Despite limitations on the functional form of alternative models that may be estimated using the Boston Fed data, due to the omission of a number of important variables, several simple changes to the model can be made to better approximate the underwriting decision process.

A number of variables that were used by the Boston Fed but that did not approach statistical significance in the sample were dropped, including the average rent/property value ratio in the census tract (RENT), the industry-specific unemployment rate (UNEMP), and self-employed status (SELFEMP). Several additional variables were incorporated into the revised models, including separate dummy variables for credit risk (NOCREDIT, DELCREDIT, BADCREDIT), mortgage payments (NOMORT, LATEMORT), and the loan-to-value ratio (LTV91, LTV96, LTV100). These variables are defined in table 3. Variables indicating consumer credit accounts 30–59 days past due and one or two late mortgage payments were not statistically significant and were dropped from the subsequent models. Similarly, the dummy variable for 80–90% LTV was not statistically significant and was excluded from subsequent models.

A dummy variable was defined to indicate whether liquid assets were sufficient to meet closing costs (LQASSET), a factor considered by lenders, as the file reviews revealed that a number of applicants were denied for this reason. A variable for educational attainment, applicant's years of school (EDUCAT), was included in the model. This variable could influence application outcome to the extent that educational attainment influences future income growth. Education also may be associated with knowledge about underwriting standards, influencing rejection rates through self-selection. The education coefficient generated by a single-equation model may reflect a combination of these factors. Education was found to be a statistically significant determinant of mortgage application outcomes by Rosenblatt (1997). Average educational attainment also had a significant influence on the number of mortgage loans observed across census tracts in the Houston area (Holmes and Horvitz 1994).

The estimation results are presented in table 5. The impact of the functional form of the model on the race effect is illustrated by comparing the revised model (the first equation presented in table 5) with the third equation in table 4 (the Boston Fed model, revised sample 1). Both models utilize the revised sample 1 as defined in table 1. The model fit associated with the revised model improves markedly, as indicated by the fall in the Akaike statistic from 635.0 to 581.6. The rank correlation between observed responses and predicted probabilities also improves, as demonstrated by the rise in Somers's D statistic from 0.79 to 0.83. Examination of the parameter estimates shows some interesting results. For example, the liquid asset variable is statistically significant, although liquid assets and net worth were not significant in any of the models specified by the Boston Fed. The liquid assets variable was dropped from their final model and therefore is not included in table 4.

The coefficient of 0.46 indicates a smaller race effect than was identified in the previous model. Comparisons between the actual and predicted denial probabilities for each group implied by the coefficient suggest that the denial rate for the sample of nonwhite applicants would fall from 30.7 to 26.6 percentage points (a difference of 4.1 percentage points) if they were treated as whites (holding other attributes constant), and the denial rate for this sample of whites would rise by from 9.2 to 11.8 percentage points (a difference of 2.6 percentage points) if they had been treated as nonwhite applicants.[16] However, the race coefficient is not statistically significant at the 0.05 level, so the hypothesis of no race effect cannot be rejected.

3.4. Verification and credit standards

The next set of regressions deals with data verification and credit standards. Unlike other factors omitted from the Boston Fed model, the Boston Fed collected information on both of these variables. Previous research suggests that both variables influence application outcomes. Zandi (1993) reported that the race effect is reduced when the Boston Fed model is supplemented by verification data and information about whether applicants met credit policy standards. Schill and Wachter (1994) and Day and Liebowitz (1994) also incorporated these variables into mortgage lending models applied to the Boston Fed data and found them to be statistically significant.

A number of potentially serious measurement problems are associated with the credit risk variables, as discussed previously. Moreover, combining data from numerous lenders may introduce aggregation bias because underwriting criteria vary across lenders. One way to avoid these problems is to include a variable indicating whether applicants met the lender's credit guidelines or standards. Munnell et al. (1992, 1996) did not include this variable in their models. Carr and Megbolugbe (1993) argue that it is inappropriate to include a credit standards variable in the model because the variable itself reflects discrimination and is used by lenders to justify denials to minority applicants. The authors contend that racial bias is the only factor that could account for the fact that nonwhite applicants were less likely to meet credit standards after controlling for other variables in the Boston Fed data set.

Although nonwhite applicants may fare worse with respect to credit standards, this difference may result from a number of reasons other than discrimination. The data in

Table 5. Revised model of application denials using alternative functional forms, revised data (sample 1), and additional credit standards and unverified data variables

Variable	Revised Model	Added Variables, Whites Only	Added Variables, All Races
RACE	0.46		0.27
	(0.0799		(0.4377)
INTERCEPT	−4.67	−0.35	−1.14
	(0.0001)	(0.8095)	(0.2895)
House	0.36	0.49	0.47
	(0.1970)	(0.3176)	(0.1942
DEBT	0.06	0.05	0.05
	(0.0001)	(0.0024)	(0.0003)
LQASSET (new)	−0.49	−0.21	−0.38
	(0.0459)	(0.6117)	(0.2204)
LTV91 (new)	0.75	−0.52	0.35
	(0.0243)	(0.4846)	(0.4348)
LTV96 (new)	4.19	5.50	4.67
	(0.0001)	(0.0001)	(0.0001)
LTV100 (new)	1.68	2.28	1.63
	(0.0005)	(0.0005)	(0.0050)
PUBREC	1.67	1.72	0.78
	(0.0001)	(0.2818)	(0.0805)
NOMORT (new)	0.34	−0.26	0.02
	(0.2470)	(0.5573)	(0.9603)
LATEMORT (new)	1.79	−0.11	0.48
	(0.0156)	(0.9232)	(0.6037)
NOCREDIT (new)	1.23	−0.39	−0.13
	(0.0181)	(0.7808)	(0.8712)
DELCREDIT (new)	1.27	−0.60	0.08
	(0.0003)	(0.4286)	(0.8686)
BADCREDIT	1.74	0.65	0.57
	(0.0001)	(0.3127)	(0.2093)
PMI	4.64	5.77	4.46
	(0.0001)	(0.0001)	(0.0001)
2–4 FAMILY	0.69	1.67	1.02
	(0.0243)	(0.0012)	(0.0085)
EDUCAT (new)	−0.06	−0.06	−0.05
	(0.1476)	(0.4074)	(0.3677)
UNVERIF (new)		2.36	2.54
		(0.0001)	(0.0001)
STANDARDS (new)		−4.64	−3.80
		(0.0001)	(0.0001)
AIC	581.6	266.3	409.8
Somers's D	0.83	0.91	0.92
Sample size	1282	1010	1282
Missing observations	15	13	18

Note: p-values are reported in parentheses.

table 2 show that nonwhite applicants experience consumer credit delinquencies almost twice as often as white applicants and have a public record of serious credit problems three times as often as white applicants. Nonwhite applicants also report less net worth, fewer liquid assets, and less income, all factors that may influence applicants' ability to meet debt obligations. Differences in meeting credit standards also may reflect many unobserved or unmeasured economic and noneconomic factors that may influence applicants' creditworthiness. Models of racial differences in credit standards are subject to the same specification flaws that affect application outcome models.

The issue relevant to modeling the lending decision is whether the credit standards variable influences the lending decision apart from any potential use of credit standards as a race ''filter.'' It is possible to test this hypothesis. A ''revised'' model that includes the credit standards (STANDARDS) and unverified data (UNVERIF) variables is estimated using only observations for white applicants. The results are provided in table 5 (Added Variables, Whites Only). Both the credit standards and unverified data variables are statistically significant. The credit risk and mortgage payment variables become statistically insignificant. Similar results (not reported here) are observed when the same model is applied solely to minority applicants. The results confirm that the two variables incorporate important information considered by lenders that is not captured by the credit history variables even when the model is estimated using only whites. The model is subsequently estimated using both nonwhite and white applicants. The results are provided in the final column of table 5 (Added Variables: All Races).[17]

The empirical results show a large and statistically significant credit standards effect and the model fit improves relative to the revised model. The unverified data variable also is statistically significant. When the two variables are included in the model, the race coefficient falls and remains statistically insignificant. The credit standards variable is primarily responsible for the change in the race estimate; race falls when the credit standards variable alone is added to the model but exhibits little change when the unverified data variable is included without the credit standards variable (results not shown).

3.5. Specifying application outcome

At present, there is no consensus as to how various types of questionable ''denials'' should be treated in the lending model. To examine the sensitivity of the race effect to different treatments, several models are estimated.

The 61 denials that did not appear to reflect lenders' willingness to provide mortgage financing are first treated as approvals, with the credit standards and unverified data variables alternately excluded and then included in the models. These questionable denials are subsequently dropped altogether from the revised sample 1 to create revised sample 2 (described in table 1), and the models (first excluding and then including the supplemental variables) are re-estimated.

The model results are provided in table 6. When the questionable outcomes are treated as approvals rather than denials (using revised sample 2), the race effect falls markedly to

Table 6. Modeling application denials, using alternative treatment of observations with outcomes that may not reflect lenders' willingness to provide mortgage financing

Variable	Outcomes Recoded as Approvals (sample 1)	Recoded and Added Variables (sample 1)	Questionable Outcomes Deleted (sample 2)	Deleted and Added Variabless (sample 2)
RACE	0.19	−0.18	0.35	0.25
	(0.4723)	(0.5710)	(0.2132)	(0.4903)
INTERCEPT	−3.89	−1.29	−4.67	−1.00
	(0.0001)	(0.1689)	(0.0001)	(0.3823)
HOUSE	0.24	0.17	0.40	0.60
	(0.3197)	(0.5962)	(0.1671)	(0.1183)
DEBT	0.04	0.03	0.06	0.05
	(0.0001)	(0.0065)	(0.0001)	(0.0004)
LQASSET	−0.64	−0.63	−0.61	−0.59
	(0.0078)	(0.0176)	(0.0184)	(0.0778)
LTV91	0.42	0.02	0.83	0.37
	(0.1868)	(0.9546)	(0.0194)	(0.4601)
LTV96	3.12	3.10	4.29	4.81
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
LTV100	1.77	1.54	2.01	2.11
	(0.0001)	(0.0036)	(0.0001)	(0.0003)
PUBREC	1.06	0.24	1.52	0.61
	(0.0009)	(0.5221)	(0.0001)	(0.1969)
NOMORT	0.20	−0.03	0.34	0.07
	(0.4848)	(0.9168)	(0.2685)	(0.8434)
LATEMORT	1.87	1.11	1.82	0.43
	(0.0091)	(0.1578)	(0.0166)	(0.6617)
NOCREDIT	1.09	−0.17	1.32	−0.11
	(0.0280)	(0.7847)	(0.0136)	(0.8916)
DELCREDIT	1.37	0.53	1.43	0.12
	(0.0001)	(0.2032)	(0.0001)	(0.8237)
BADCREDIT	1.97	1.23	1.97	0.81
	(0.0001)	(0.0008)	(0.0001)	(0.0929)
2–4Family	0.42	0.47	0.69	1.00
	(0.1610)	(0.1630)	(0.0320)	(0.0146)
EDUCAT	−0.06	−0.04	−0.07	−0.05
	(0.1186)	(0.3679)	(0.1094)	(0.3637)
UNVERIF		1.26		2.60
		(0.0006)		(0.0001)
STANDARDS		−2.47		−3.97
		(0.0001)		(0.0001)
AIC	594.6	502.2	531.3	361.8
Somers's *D*	0.74	0.85	0.78	0.91
Sample size	1282	1282	1221	1221
Missing observations	15	18	15	18

Note: p-values are reported in parentheses.

0.19. The race effect becomes negative (−0.18) when the credit standards and unverified data variables are included. Neither race effect is statistically significant.

After deleting observations for which the outcome variable may not reflect lenders' willingness to provide mortgage financing (revised sample 2), a statistically insignificant race effect of 0.35 is generated by the model (Questionable Outcomes Deleted, table 6). The race effect declines to 0.25 when the two additional variables (meets credit standards, unverified information) are added to the model (Deleted and Added Variables, table 6). None of the parameter estimates for race in table 6 is statistically significant. The model fit improves dramatically and the predictive ability increases when the questionable denials are deleted, as compared to when these observations are recoded (Deleted and Added Variables, sample 2 vs. Recoded and Added Variables, sample 2).

3.6. Aggregation issues

The data consist of mortgage applications submitted to numerous lenders in the Boston area. Estimates based on combined data may obscure important differences in lending practices among institutions. Of the 70 FDIC-supervised lenders in the sample, 47 (67%) did not reject a single nonwhite applicant. Of these lenders, 23 reported no applications from nonwhites and another 24 approved a total of 69 applications from nonwhites.

Closer examination revealed that 47% of all denials to nonwhites in the sample of FDIC-supervised institutions (revised sample 1) were accounted for by one lender. Furthermore, denial rates were extraordinarily high at this institution, which denied one out of every two applications from nonwhites. If high denial rates for nonwhites reflect lending bias, then one might expect racial discrimination to be rampant at this institution. However, closer inspection does not appear to support this contention. The institution, which is the sole minority-owned bank in the Boston area, conducted extensive minority outreach programs and participated actively in a number of affordable housing programs. Its commitment to lending to minority applicants is demonstrated by the fact that 82% of all mortgage loans approved by this institution were to nonwhite applicants. It is difficult to reconcile the fact that this lender accounts for such a large proportion of minority denials with allegations that (1) high denial rates for nonwhites are evidence of discrimination and (2) lending bias in the Boston area is widespread.[18]

Munnell et al. (1996) estimated separate models for institutions reporting a large volume of loans to nonwhites (active lenders) and those making few mortgage loans to nonwhites (inactive lenders). The race effect was greater for active lenders to nonwhites, implying a greater degree of discrimination by these lenders. The authors claim that the statistical significance of race in both models demonstrates that the influence of race "appears to be pervasive in the market" (p. 41). An alternative interpretation of these results, given the intuitively plausible hypothesis that lenders most active in nonwhite communities are less likely to engage in racial discrimination, is that factors omitted from these models are responsible for generating the observed race effects.

4. Conclusions

Several prior studies of mortgage application data identified statistically significant race effects, indicating that minority applicants are more likely to be denied after accounting for a number of other factors. The authors of these studies argue that this race effect is attributable to racial discrimination by mortgage lenders because they have controlled for all other important factors. The research presented here calls into question the conclusion that disparities in denial rates are caused by racial discrimination. This paper addresses a relatively narrow issue: Do application outcome models represent the process by which lenders evaluate mortgage applications? Several major problems with these models are examined. First, the functional form of the models do not capture the complex interactions between variables and the nonlinear weights associated with many variables. Second, many of the variables included in the models are crude proxies for the factors of interest to lenders. Third, a number of factors considered by lenders are omitted altogether. Fourth, the application outcome variable does not represent lenders' willingness to provide mortgage financing. The theoretical framework of the lending decision was examined to illustrate the potential for bias that may arise as a consequence of inadequacies with the simple models that have been applied to mortgage lending.

The empirical results presented in this paper demonstrate that estimates of the race effect are quite sensitive to variations in model specification, despite limitations with the quantity and quality of the Boston Fed data. The credit risk variables used by the Boston Fed are a poor proxy for lenders' assessment of credit risk; the credit standards variable clearly reflects information that is not captured by the other credit risk variables. The credit standards presumably incorporates more of the information considered by lenders but could itself be influenced by the race of the applicant, thereby introducing ''included variable bias.''[19] The empirical results suggest that a potential bias results from using a poor proxy for credit risk, but the use of the lender's assessment may bias the race effect in the other direction. The magnitude of these sources of bias cannot be measured precisely using the data available. Given the importance of credit risk in the lending decision, a better measure of credit risk is needed. Unverified data represent another problem. The empirical results suggest that the inability to verify application information has a negative impact on the likelihood of obtaining a loan, but this variable also was omitted from the Boston Fed model. Future studies should attempt to collect verified data that lenders use to evaluate applicants' creditworthiness.

The statistical significance of the variable indicating whether applicants had sufficient financial resources to satisfy closing costs suggests that lenders consider this factor, although this factor was not included in prior analyses of the Boston Fed data. The view that lenders consider whether applicants have sufficient assets is reinforced by the fact that denials based on insufficient assets were found in the file reviews. However, the liquid asset variable included in the models reported in tables 5 and 6 is a poor proxy in several respects. First, no information on actual closing costs was available. In addition, applicants rely on gifts from relatives, equity available to current homeowners, and a variety of financial resources other than liquid assets to finance closing costs.

A number of other problems with application outcome models have been discussed in the literature. Single-equation models provide reduced-form estimates of a multiequation system that correspond to various sequential decisions by potential applicants and lenders. Many important factors that influence this process are unobserved. Such models cannot determine the extent to which differences in denial rates across racial groups are influenced by information asymmetries, preapplication screening, community outreach programs, or the ability of applicants to modify applications to overcome obstacles raised by lenders. Even if the application outcome models include all the factors considered by lenders, the potential for self-selection or simultaneous-equations bias cannot be dismissed.

Although the revised models presented in this study represent improvements over the original Boston Fed model, none of the estimated race effects generated by these models can be considered precise given the limitations of the data. The instability of the race estimates suggest that skepticism about the magnitude, statistical significance, and even the sign of the race effect is warranted. Nevertheless, the statistical results presented in tables 5 and 6 suggest that better specification of the credit risk variable and more precise coding of the outcome variable may explain the racial disparities in denial rates. Given the problems inherent with estimating simple, reduced-form loan application models, the empirical results should be interpreted with caution, particularly given the inclination to use these models for such purposes as monitoring institutions and targeting lenders for enforcement actions.

Appendix table. Mean values of regression variables for revised sample

Variable	Mean	Standard Deviation
APPROVE	0.86	0.34
RACE	0.21	0.42
HOUSE	0.23	0.42
DEBT	33.64	10.66
NTWORTH ($000)	$234	1237
LQASSET	0.63	0.48
LTV	0.76	0.18
LTV91	0.10	0.30
LTV96	0.02	0.13
LTV100	0.03	0.17
PUBREC	0.07	0.25
NOMORT	0.70	0.46
LATEMORT	0.01	0.10
MORTPAY	1.76	0.53
NOCREDIT	0.04	0.19
DELCREDIT	0.07	0.26
BADCREDIT	0.08	0.27
CREDIT	2.15	1.64
PMI	0.04	0.20
RENT	0.08	0.17
2-4FAMILY	0.13	0.33
SELFEMP	0.12	0.32
UNEMP	3.84	2.08
EDUCAT	15.19	2.95
UNVERIF	0.06	0.24
STANDARDS	0.90	0.31

Acknowledgments

The views expressed in this paper are those of the author and should not be attributed to the FDIC. The author is grateful to Steven Guggenmos for exceptional research assistance and to George Benston, Lisa Daniel, James Freund, Anthony Yezer, and an anonymous referee for helpful comments and suggestions.

Notes

A preliminary and unfinished draft of this paper was provided to the authors of the Boston Fed Study, at their request, on the condition that the work not be quoted or cited. Munnell et al. (1996) subsequently included an extensive critique of the document in the version of their paper published in the *American Economic Review*. However, because their comments refer to the preliminary draft, which utilized incomplete data, many of their assertions are factually incorrect. A more complete response to their critique is addressed in a reply submitted to the *American Economic Review*.

1. A Gallup survey released by the Mortgage Bankers Association of America (March 9, 1994) documents the public's perception of lending bias. Among individuals who have never applied for a mortgage, 70% of blacks, 49% of Hispanics, and 49% of whites responded that they believed that discrimination occurs in mortgage lending.
2. The model specifications and regression results reported in their working paper differ slightly from those published in the *AER*. In the latter paper, Munnell et al. exclude rent/value in tract and substitute three dummy variables for a continuous loan/appraised value. The authors find that ''the probability of denial increases 8.2 percentage points for a minority applicant'' (p. 33). Similar race effects are reported for a number of other model specifications.
3. The author accompanied examiners at a number of institutions.
4. In contrast, Day and Liebowitz (1994) found that the race effect appeared to be quite sensitive to the inclusion of a small number of influential observations. Subsequent work by Stengel and Glennon (1995), using mortgage application data from three national banks, showed that the race effects were sensitive to variations in the function form of the model.
5. An applicant with a housing expense ratio of 14%, for example, was reported as 0.14 although the survey instructed lenders to report the percentage in whole numbers (i.e., as 14).
6. In some cases liabilities appeared excessively high even though assets and income appeared to be normal. For example, 30 applicants with total liabilities exceeding $500,000 (up to a maximum of $8.2 million) reported monthly debt obligations of $1000 per month or less; half of these reported monthly debt obligations below $100. Liabilities exceeded total assets in 15 cases. However, 26 of the 30 were approved for loans.
7. There is some justification for excluding applications with higher loan-to-value ratios, but a 30% ratio was selected to avoid further reducing the sample size. The models were estimated excluding applications with less than a 50% loan-to-value ratio to examine the sensitivity of the parameter estimates. The empirical results were relatively insensitive to the choice of the loan-to-value ratio within this range, although the race coefficients were slightly smaller when these observations were deleted.
8. The file reviews identified one case where a lender had provided a written commitment to provide mortgage financing but title problems ultimately thwarted the sale of property. Although a written commitment could be interpreted as an indication that the lender was willing to provide financing to the applicant in the absence of such problems, under HMDA this outcome is reported as a lender denial.
9. Under HMDA, lenders are required to report applicant's race or ethnicity; in this sample, applicants are characterized as either black, white, or Hispanic. This approach produces results that are somewhat ambiguous because race and ethnicity are not mutually exclusive: Hispanic applicants also may be white or black. The Boston Fed study compared denial rates of Hispanics and blacks (both classified as nonwhite applicants) with whites.

10. The mean values showed larger differences between nonwhite and white applicants. The mean value of liquid assets for whites, \$102,400, is more than three times that for nonwhite applicants (\$32,600). Similarly the mean net worth of whites, \$275,000, is over three times the \$82,700 that reported by nonwhite applicants.
11. Benston, in the October 1979 version of his review of redlining research, noted a study of mortgage lending in which observations with missing variables were omitted. Because the banks sought to provide complete documentation for black applicants who were denied but were less meticulous with respect to applications from whites, this procedure primarily excluded denied white applicants, thereby biasing the statistical results.
12. According to the Survey of Consumer Finances, the median net worth of white families was \$58,500 in 1989 vs. \$4,000 for nonwhite and Hispanic families. The mean net worth was \$203,800 and \$45,900, respectively (Kennickell and Shack-Marquez 1992). In addition, survey data from the National Association of Realtors (1992) show that ''whites were twice as likely as blacks to use equity from a previously owned home.''
13. Munnell et al. report additional model specifications that include the separate dummy variables for mortgage payment and consumer credit.
14. The denial rates reported in this section are the same as those provided in table 2 for the original data and for the revised sample 1. However, the mean denial probabilities calculated from each subsequent model differ slightly from these denial rates because the means are derived from slightly different samples; a small number of observations with missing data are excluded from each of the logit regressions. The mean denial rates also vary somewhat as additional applications are deleted to produce revised samples 2 and 3.
15. The AIC statistic is discussed in Amemiya (1981), p. 1505, and SAS *Users Guide* (1990), vol. 2, p. 1088.
16. The magnitude of the race effects discussed previously are derived from models estimated on nonwhite and white applicants together. This approach constrains the parameter estimates to be the same for nonwhite and white applicants. As a result, the different race effects for nonwhite and white applicants result solely from differences in attributes. Constraining the parameter estimates may be unjustified if lenders emphasize different factors for nonwhite and white applicants. To evaluate the extent to which the estimated race effect is sensitive to the use of a single equation, the revised model in table 5 was estimated separately for nonwhite and white applicants. The parameter estimates from the model of nonwhites were used to generate denial probabilities for white applicants and the parameter estimates from the model for whites were used to generate denial probabilities for nonwhite applicants. Similar techniques have been applied to estimate race effects in earnings equations (e.g., Blinder 1973, Malkiel and Malkiel 1973, Oaxaca 1973) and to mortgage lending as well (e.g., Benston, Horsky and Weingartner 1978, Benston and Horsky 1979). Munnell et al. (1992, 1996) also estimate race-specific models.

    The parameter estimates from the race-specific equations imply that minority status would generate a denial rate increase of 3.7 percentage points for nonwhite applicants and 2.9 percentage points for whites. These race effects are quite similar to the effects estimated using the single-race model.
17. Day and Liebowitz (1994) presented a similar analysis (examining the impact of the credit standard variable for white and nonwhite applicants separately) applied to the full Boston Fed data set and obtained similar results.
18. The statistically significance of the race effect generated by applying the Boston Fed model specification to the revised data (as shown in the last column of table 4) is eliminated when applications reported by this one lender are excluded from the analysis. The race coefficient falls from 0.67 to 0.35, and the p-value rises from .0058 to .2048. The race effects generated by all subsequent models also are statistically insignificant. The results, not provided in this paper, are available from the author on request.
19. This concept is discussed in Killingsworth (1993).

References

Amemiya, Takeshi. (1981). ''Qualitative Response Models: A Survey,'' *Journal of Economic Literature* 99 (December), 1483–1536.

Barth, James R., Joseph J. Cordes, and Anthony M. Yezer. (1979). ''Financial Institution Regulations, Redlining and Mortgage Markets,'' in *The Regulation of Financial Institutions*, pp. 93–107 Conference Series No. 21, Federal Reserve Bank of Boston and the National Science Foundation.

Benston, George J. (1981). ''Mortgage Redlining Research: A Review and Critical Analysis,'' *Journal of Bank Research* 12 (Spring), 8–23; reprinted in *The Regulation of Financial Institutions*, pp. 144–195, Conference Series No. 21, Federal Reserve Bank of Boston and the National Science Foundation (October 1979).

Benston, George J., and Dan Horsky. (1979). ''Redlining and the Demand for Mortgages in the Central City and Suburbs,'' *Journal of Bank Research* 10 (Summer), 72–87.

Benston, George J., and Dan Horsky. (1992). ''The Relationship Between the Demand and Supply of Home Financing and Neighborhood Characteristics: An Empirical Study of Mortgage Redlining,'' *Journal of Financial Services Research* 5 (February), 235–260.

Benston, George J., Dan Horsky, and H. Martin Weingartner. (1978). ''An Empirical Study of Mortgage Redlining,'' Monograph 1978-5, Monograph Series in Finance and Economics. New York: New York University Graduate School of Business Administration, Solomon Brothers Center for the Study of Financial Institutions.

Blinder, Alan S. (1973). ''Wage Discrimination: Reduced Form and Structural Estimates,'' *Journal of Human Resources* 8(4), 436–455.

Carr, James H., and Isaac F. Megbolugbe. (1993). ''The Federal Reserve Bank of Boston Study on Mortgage Lending Revisited,'' *Journal of Housing Research* 4(2), 277–313.

Day, Ted, and S. J. Liebowitz. (1994). ''Mortgages, Minorities, and Discrimination,'' draft manuscript.

Glennon, Dennis, and Mitchell Stengel. (1994). 'An Evaluation of the Federal Reserve Bank of Boston's Study of Racial Discrimination in Mortgage Lending,'' Working Paper 94-2, Office of the Comptroller of the Currency, Washington DC.

Holmes, Andrew, and Paul Horvitz. (1994). ''Mortgage Redlining: Race, Risk, and Demand,'' *Journal of Finance* (March), 81–99.

Horne, David K. (1994). ''Evaluating the Role of Race in Mortgage Lending,'' *FDIC Banking Review* (Spring–Summer), 1–15.

Hunter, William C., and Mary Beth Walker. (1996). The Cultural Affinity Hypothesis and Mortgage Lending Decisions,'' *Journal of Real Estate Finance and Economics*, 13(1), 57–70.

Kennickell, Arthur, and Janice Shack-Marquez. (1992). ''Changes in Family Finances from 1983 to 1989: Evidence From the Survey of Consumer Finances,'' *Federal Reserve Bulletin* 78 (January), 1–18.

Killingsworth, Mark R. (1993). ''Analyzing Employment Discrimination: From the Seminar Room to the Courtroom,'' *American Economic Review* 83 (May), 67–72.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Economics*. New York: Cambridge University Press.

Maddala, G. S., and Robert P. Trost. (1982). ''On Measuring Discrimination in Loan Markets,'' *Housing Finance Review* 1 (July), 245–68.

Malkiel, Burton G., and Judith A. Malkiel. (1973) ''Male-Female Pay Differentials in Professional Employment,'' *American Economic Review* 63 (September), 693–705.

Mortgage Bankers Association of America. (1994). ''Gallup Survey Shows People Perceive Many Barriers To Owning A Home,'' news release and executive summary, (March 9).

Munnell, Alicia H., Lynn E. Browne, James McEneaney, and Geoffrey M. B. Tootell. (1992). ''Mortgage Lending in Boston: Interpreting HMDA Data,'' Working Paper 92-7, Federal Reserve Bank of Boston.

Munnell, Alicia H., Geoffrey M. B. Tootell, Lynn E. Browne, and James McEneaney. (1996). ''Mortgage Lending in Boston: Interpreting HMDA Data,'' *American Economic Review* 86 (March), 25–53.

National Association of Realtors. (1992). *Survey of Homeowners and Renters: Key Findings*. Washington, DC: National Association of Realtors.

Oaxaca, Ronald. (1973). ''Male-Female Wage Differentials in Urban Labor Markets,'' *International Economic Review* 14 (October), 693–709.

Rachlis, Mitchell B., and Anthony M. J. Yezer. (1993). ''Serious Flaws in Statistical Tests for Discrimination in Mortgage Markets,'' *Journal of Housing Research* 4(2), 315–336.

Rosenblatt, Eric. (1997). ''A Reconsideration of Discrimination in Mortgage Underwriting with Data from a National Mortgage Bank,'' *Journal of Financial Services Research* 11(1,2), 109–131.

SAS Institute Inc. (1990). *SAS/STAT User's Guide*, Version 6, 2, 4th ed. Cary, NC: SAS Institute.

Schill, Michael H., and Susan M. Wachter. (1993). ‘‘Racial and Ethnic Geographic Disparities in Home Mortgage Lending,’’ *Journal of Housing Research* 4(2), 245–275.

Schill, Michael H., and Susan M. Wachter. (1994). ‘‘Borrower and Neighborhood Racial and Income Characteristics and Financial Institution Mortgage Application Screening, *Journal of Real Estate Finance and Housing* 9, 223–239.

Stengel, Mitchell, and Dennis Glennon. (1995). ‘‘Evaluating Statistical Models of Mortgage Lending Discrimination: A Bank-Specific Analysis,’’ Working Paper 95-3, Office of the Comptroller of the Currency, Washington DC.

Yezer, Anthony, Robert Phillips, and Robert Trost. (1994). ‘‘Bias in Estimates of Discrimination and Default in Mortgage Lending: The Effects of Simultaneity and Self Selection,’’ *Journal of Real Estate Finance and Housing* 9, 197–215.

Zandi, Mark. (1993). ‘‘Boston Fed’s Bias Study Was Deeply Flawed,’’ *American Banker* (August 19).