Mortgage lending in Boston: a reconsideration of the evidence

Author: Glenn W. Harrison
Date: Jan. 1, 1998
From: Economic Inquiry(Vol. 36, Issue 1)
Publisher: Blackwell Publishers Ltd.
Document Type: Article
Length: 4,673 words

Abstract:
Is there statistical evidence of racial discrimination in home mortgage markets? The Boston Fed recently addressed this concern head-on, by collecting all available data from loan applications in Boston. They find that the extent of discrimination is reduced after one accounts for all of the confounding variables measured in these applications, but that it remains statistically significant. However, their strong conclusions are unwarranted due to the use of invalid statistical methods. Correctly evaluated, their data provide no significant evidence of racial discrimination in mortgage markets.

Full Text:
I. INTRODUCTION

Discrimination in the marketplace is often measured using statistical methods. In order for the results to be interpretable as some form of discrimination it is necessary that these methods be firmly based on economic theories as to how these markets function. The reason for this is that it is easy to confuse voluntary acts of segregation from involuntary acts of discrimination by the simple examination of statistical entrails. The former is presumably acceptable to society, whereas the latter is not. Any society that values its freedoms will want to ensure that a mis-classification of the former as evidence of the latter is avoided.

Claims of racial discrimination in mortgage lending provide an excellent arena to illustrate the issues that arise in statistically measuring discrimination. In the United States data on mortgage applications has been collected for several years under the Home Mortgage Disclosure Act (HMDA). These data show a persistent tendency for blacks to be denied mortgages at a greater rate than whites, leading many political commentators to claim that they prove the existence and severity of racial discrimination.

One problem with the statistical measurement of discrimination is that the available data are often inadequate. The HMDA data has been criticized for not collecting all of the information germane to the loan decision, such as the credit history or personal characteristics of the loan applicant and co-applicant. The argument is that blacks are denied loans more than whites because of these characteristics, rather than the color of their skin. Whether or not this, in turn, reflects prior or concurrent racial discrimination in other markets (e.g., in employment or housing or education) is an intriguing question with important implications for policy reform in mortgage lending regulations.(1)

In an effort to address these concerns, the Federal Reserve Bank of Boston undertook to augment the basic HMDA data in the Boston area with ancillary data on loan applicants: see Munnell et al. [1992, 1996]. The vast bulk of these extra data were collected from the original loan application file at the bank of original application, indicating that the information was viewed by the bank as relevant to the loan decision.

The Boston Fed study found that allowing for the characteristics of black applicants reduced the statistical importance of the applicant's race, but that some racial discrimination remained. Specifically, they find that black and Hispanic applicants are roughly 60% more likely to be turned down just because of their ethnic origin. The denial rate in this sample is about 17% for such applicants, rather than the 11% denial rate for "statistically equivalent" whites. Although the study has been the subject of some controversy in the popular press(2) and policy circles,(3) it remains an important effort to address these issues by directly collecting the primary data.

There are, however, some problems with the way in which these data are interpreted by Munnell et al. [1992, 1996]. These problem concern their statistical methods, even if one accepts the economic model of mortgage lending decisions underlying their study. Simple corrections to their statistical procedures removes any evidence of discrimination. An important conclusion, echoing the Boston Fed study itself, is to point out the dangers of drawing inferences about discrimination based on incomplete data.

Moreover, the simple statistical corrections considered here are the standard fare of textbooks and not from the fancy frontiers of

econometrics. Hence it is likely to be useful in general to see why those textbook lessons are there by having them demonstrated explicitly in an important policy setting.

II. REPLICATING THE BOSTON FED RESULTS

The Boston Fed study began with the existing HMDA database for mortgage applications in the Boston area for 1990.(4) The Fed requested that banks augment their HMDA reports for all applications made by blacks and Hispanics, and by a random sample of applications made by whites. The result was a sample of 722 black and Hispanic applicants and 2340 white applicants.

The original HMDA reports only contained information on one directly economic variable: applicant income. Information was requested on an additional 38 variables deemed likely to influence the approval decision, and indeed likely to be found in the original loan application file. These additional variables included information on the ability of applicants to support the loan, the reliability of the borrower against risk of default as measured by credit history and income stability, potential default loss to the primary lender, characteristics of the loan, and personal characteristics.

The economic model used in the Boston Fed study(5) focuses solely on the mortgage approval process. They explicitly assume (p. 27) that individual banks take the market interest rate as fixed, and simply decide whether the loan would result in an unacceptably high risk of default. This risk is mitigated in the short-run by the ability of banks to sell the loan on a secondary market, but persistent defaults of such loans will damage the bank's ability to make future sales on the secondary market. Hence the primary concern of the bank is to determine the likelihood of default.

In effect, then, the Boston Fed study uses the standard model in which the approval or denial decision is related to the characteristics noted above. These characteristics are presumed to determine the minimum interest rate which the bank would require in order for the loan to increase its expected profit at the margin. If this minimum rate exceeds the going market interest rate then the bank is assumed to deny the loan; otherwise the loan is approved.(6)

The statistical analysis of the Boston Fed study revolves around one basic equation. The dependent variable to be explained is binary: approval is coded as a 0 and denial as a 1 for each individual.(7) The explanatory variables used in their "preferred equation" (their Table 2, p. 34) are defined in Table I and listed in Table II. The Boston Fed study uses a subset of the variables available from their data set: the additional variables used later are listed in Table III. Boston Fed Table 3 (p. 36/37) contains the results of adding several of these to the preferred equation, one group at a time. In each case they find no substantial change in the coefficient on the race variable (BORH) or in its statistical significance.

The results in Table II can be easily interpreted.(8) The first panel shows some simple descriptive statistics, to orient one to the sample. Overall 14.534% of mortgage applications in the sample were denied, and that the sample consisted of 23.371% blacks and Hispanics. The next panel shows the estimation results, using the logit specification for the basic equation following the Boston Fed study. We see that many of the explanatory variables are individually significant (i.e., a low probability of the t statistic being equal to zero by chance alone), and that the overall equation is statistically significant (i.e., we can easily reject the hypothesis that all coefficients are jointly non-zero by chance alone).

The key coefficient on BORH is also statistically significant. The coefficient value is 0.70776 with a standard error of 0.1385. By itself this coefficient does not mean much in terms of the effect of race on the probability of mortgage denial, since this parametric specification is not additive in the variables.

To determine the effect on the probability of denial requires additional calculations.(9) For the marginal effect of a continuous variable these marginal effects may be calculated readily, and are shown at the bottom of Table II. For the marginal effect of a binary variable, such as BORH, one can compute these marginal effects but care is needed in interpreting them. They refer to a change in the probability that any given individual is classified BORH; at the individual level this may not make much sense, but at an aggregate level it is easy enough to interpret.(10)

A given change in the proportion of the sample classified as BORH would result in an increase in mortgage denial of 6.361%. This means that a subject in this sample that had all of the average characteristics of the sample, which is not the same as the average characteristics of blacks and Hispanics in the sample, would be denied a mortgage 6.3% more often if he were also BORH compared to not being BORH. The standard error on this estimated marginal effect is 1.226%, implying that there is a statistically significant effect of being BORH on the probability of mortgage denial.(11)

How do these results compare in detail to those in the Boston Fed study? Comparing these coefficients to those in the Boston Fed Table 2 (p. 34) we find that all coefficient values are close but not identical. There are several simple reasons for this slight disparity. The public use database is smaller than the one used in this study, by about 130 observations. This is only 4.2% of the original sample, and presumably represents subjects deleted to preserve confidentiality requirements. Second, one variable in the Boston Fed equation, RVAL, was not provided on the public use database and was therefore not used in our estimates. Third, we may have classified some variables differently than the Boston Fed study. For example, there are some minor ambiguities in how the dependent variable is classified, which could cause minor disparities in results. Finally, there could be small differences in the estimation results from different statistical packages, although for a common model specification such as logit this is relatively unlikely. In any event, these estimates are extremely close to those of the Boston Fed study, supporting the conclusion that their results have been adequately replicated as the basis for further evaluation.

III. SOME SIMPLE STATISTICAL CORRECTIONS

Re-weighting the Samples

The sample of blacks and Hispanics in the database is selected in a non-random and exhaustive manner, for reasons that are understandable given their small size (p. 30). Although this causes no bias in the estimation procedure,(12) it can cause some differences in the computed marginal effects which are evaluated at sample means. In the Boston Fed study this difference in stratification is substantial: they estimate the percentage of blacks and Hispanics in the Boston PMSA at 11.3%, but their final sample contains 23.6% blacks and Hispanics.(13)

If we regenerate the marginal effects from the Boston Fed specification using appropriate sample weights for the observations with respect to racial composition, the marginal effect of being black or Hispanic on mortgage denial drops from 6.36% to 5.1%. The standard error increases slightly from 1.23% to 1.9%, so this is still a statistically significant effect. However, it is somewhat smaller in absolute magnitude than the original estimate. In all subsequent analysis we use these weighted results, unless otherwise noted.

The Kitchen Sink Effect

An obvious issue that arises in evaluating the Boston Fed results is why they choose to include some explanatory variables in their main estimating equation and not others. It is difficult a priori to justify the exclusion of the variables listed in Table III. Undoubtedly the customary practice of tinkering with the specification has been followed, such that what is reported as the "preferred equation" is the result of ex post model selection. The natural fear in this case is that the preferred equation does not adequately represent the set of inferences that are possible with the data set and a different set of priors as to which variables "ought" to be included in the final equation (Learner [1978]).

To some extent the Boston Fed study attempts to address this concern in the their Table 3 (p. 36-37)(14) by examining a wide variety of alternative specifications. In all cases they find that the coefficient on the race variable changes negligibly, typically in the second significant digit. These specification searches provide an explicit clue as to the set of variables that were accorded some relevance by the Boston Fed study, even if they did not survive to be reported in the preferred equation. Moreover, they repeatedly note that their survey attempted to collect all variables deemed important in the lending decision: "... any empirical study of mortgage lending must include those variables that lenders actually consider when making their decisions, rather than simply what they ought to consider." (p. 27).

A difficulty with the way in which these alternatives was examined is that they entailed partial modifications of the preferred equation. To take a contrived example, it could be that adding information about the sex of the applicant makes little difference to the effect of race, but adding information about the sex of the applicant and the sex of the co-applicant does. One simple way to check for this type of interaction effect, which is inevitable when the explanatory variables exhibit any degree of multicollinearity, is to include all variables that are deemed relevant.

The resulting "kitchen sink" equation will have coefficient estimates that are unbiased, even if some of these variables are "irrelevant" (see Greene [1991, 261-262]). If the variables are indeed irrelevant, then the worst that can happen is that the standard errors on the coefficient estimates will be inflated. This could be expected to result in the coefficient on BORH remaining positive, but becoming statistically insignificant. However, this is why it is important to recognize that the Boston Fed study, in its Table 3, does indicate that it views these variables as being "relevant." Hence we may include them without fear that they can be deemed to be "irrelevant" from the perspective of the priors employed in that study. The use of quotation marks in the preceding sentences should alert one to the fact that we are referring to subjective priors here, and that the only concern is to evaluate the database relative to the revealed priors of the Boston Fed study.

The results of the kitchen sink specification, which includes virtually all of the variables available in the public use database, are shown in Tables IV and V. The results are astonishing: race is no longer a significant variable! The specification now includes two race variables, one for the applicant as before (BORH) and one for the co-applicant (COBORH). Neither is significantly different [TABULAR DATA FOR TABLE IV OMITTED] from zero at any standard critical level, nor are they jointly different from zero using a Wald test (see Greene [1991, 128-130]).(15)

The marginal effect of the race of the applicant on the mortgage decision is now only 1.26%, with a standard error of 2.3%. Excluding the race of the co-applicant, which is positively correlated with the race of the applicant, makes no difference to these qualitative conclusions.

An even stronger statement can be made by evaluating every possible extension of the core Boston Fed specification to include extra variables found in the application file. Random evaluation of over 10,000 different combinations of these extra variables resulted in an estimated marginal effect that was never significant at the 5% level, and only significant at the 10% level in one case. Hence our claim is robust to a much wider range of prior beliefs than implicitly used in the Boston Fed.

I conclude from this exercise that the Boston Fed was correct to emphasize the dangers of drawing conclusions from incomplete data, but that it has fallen into essentially the same trap as the result of piecemeal attempts to test its basic statistical specification.

Excluding Hispanics

Hispanics are notoriously difficult to classify as a racial and ethnic group. For example, the Boston Fed study notes that "... 51 applications that a suburban bank had coded as Hispanic in its original HMDA submission were found to be white" [1992, 20] when that bank re-evaluated their originally submitted data. The problems are far less severe with respect to blacks. Hence it could be that the tendency to classify an applicant as Hispanic could be correlated with certain attributes that the applicant reveals during the application process, rather than intrinsic racial status. If these attributes are also correlated with the likelihood of the loan being approved, there could be a serious bias in the extent of the estimated discrimination.

The original Boston Fed specification can be regenerated using a simpler race variable, BLACK, rather than BORH. Replicating that

specification without any sample weights, we estimate a marginal effect of being black on mortgage denial of 5.58%, which is also statistically significant as with BORH. However, when we simply re-weight these data this marginal effect drops to 4.75% with a standard error of 3.3%, making it significant only at the critical level of 15%. Using traditional levels of significance we would conclude that race is not a statistically significant variable in this case.

When we adopt the preferred kitchen sink specification, we arrive at essentially the same conclusions as when we used the broader race variable BORH: race, either of the applicant or co-applicant, is insignificant.

IV. POLICY IMPLICATIONS

The Boston Fed is to be praised for directly addressing some of the criticisms of the casual policy conclusions drawn from the HMDA data. Only additional data, of the kind they collected, will allow the reasoned resolution of endless debates between uninformed zealots. A re-examination of their preferred statistical approach leads to the simple conclusion that no evidence for discrimination exists in these mortgage markets when all of the relevant variables are included.

Whether discrimination occurs in other markets that determine the values of "all of the relevant variables" remains an open question. If it does, and society wants to eradicate it with intervention, then second-best alarms should go off at attempts to intervene in mortgage markets to correct problems arising in other markets. Mitigation of some of the effects of discrimination can all too easily be confused with eradication of the root causes.

TABLE I

Definitions of Core Variables

DECISION Denial. If the loan is originated, approved but not accepted by the applicant, or purchased by the institution, code as 0; if it is denied, withdrawn or the file closed for incompleteness, code as 1.

OBRAT Housing Expense/Income. Code as 1 if ratio exceeds 0.3, since 0.28 is a secondary market guideline (p. 3); 0 otherwise.

TOBRAT Total Debt Payments/Income. Ratio of total obligations to total income.

TOBRAT2 Total Debt Payments/Income. Code as I if ratio exceeds 0.36, since that is a secondary market guideline (p. 3); 0 otherwise.

NETW New Wealth. Total assets minus total liabilities, in thousands. The co-applicant's reports were used is separate statements were completed.

CONSPAY Consumer Credit History. Code as 1 if no "slow pay" accounts; 2 if one or two slow pay accounts 3 if more than two slow pay accounts; 4 is insufficient history for determination; 5 is delinquent history with 60 days past due; and 6 if serious delinquencies with 90 days past due.

MORTPAY Mortgage Credit History. Code as 1 if no late payments; 2 is no payment history; 3 if one or more late payments; and 4 if more than two late payments.

PUBREC Public Record History. Coded as 1 if there is any public record of credit problems; 0 otherwise.

URIA Probability of Unemployment. State unemployment rate for applicant's industry in 1989.

SELF Self-Employed. Coded as 1 if applicant is self-employed; 0 otherwise.

LAV Loan/Appraised Value. Value of the loan amount requested divided by appraised value.

INSGET Denied Private Mortgage Insurance check.

RVAL Rent/Value in Tract. "Rental income divided by estimate of value of rental property from Census." (p. 28) This variable was not in the public use data set of 9/28/93 provided by the Boston Fed.

FAMILY Purchasing Two- to Four-Family Home. If purchasing. a single-family home or a condo, code as 0; code as 1 if purchasing a two to four-family home.

BORH Race. If the applicant is black or Hispanic, code as 1; 0 otherwise.

TABLE III

Definitions of Additional Variables

TOBRAT2 Total Debt Payments/Income. Code as 1 if ratio exceeds 0.36, since that is a secondary market guideline (p. 3); 0 otherwise.

COBORH Co-Applicant Race. If the co-applicant is black or Hispanic, code as 1; 0 otherwise.

BLACK Black. If the applicant is black, code as 1; 0 otherwise.

COBLACK Co-applicant Black. If the co-applicant is black, code as 1; 0 otherwise.

WHITE White. If the applicant is white, code as 1; 0 otherwise.

COWHITE Co-applicant White. If the co-applicant is white, code as 1; 0 otherwise.

SEX Sex. Code as 1 for a female; 0 otherwise.

COSEX Co-applicant Sex. Code as 1 for a female; 0 otherwise.

MARRIED Marital Status.

NDEP Number of Dependents.

OLD Age. If the applicant's age exceeds the median age for the Metropolitan Statistical Area (MSA) of the property, code as 1; 0 otherwise.

GIFT Gift in Down Payment. Code as 1 if gift or grant is part of the down payment; 0 otherwise.

COSIGNER Co-signer of application. Code as 1 if there was a co-signer; 0 otherwise.

UNVER Unverifiable Information. Code as 1 if some information on the application was unverifiable; 0 if all information was verifiable.

NREVIEW Number of Reviews. Number of times the application was reviewed by the underwriter before the final loan decision was made.

I am grateful to McKinley Blackburn, Kevin Gray and David Kennison for helpful comments, and to James McEneaney of the Federal Reserve Bank of Boston for sharing data. None of the views here should be attributed to my employers, research sponsors, or overlapping generations.

1. The observed data may also reflect rational attempts by firms or banks to minimize any risk of tripping some statistical wire that might set off a regulatory bomb, as discussed by Collet, Harrison and Rutherford [1996].

2. For example, see Liebowitz [1993] and Browne [1993].

3. Cited in Munnell et al. [1996]. Many of these concerns have to do with alleged data inconsistencies. Some of these claims are reviewed in an appendix, available on the web site http://theweb.badm.sc.edu.proxy.lib.duke.edu/glenn/hmda.htm.

4. The HMDA legislation was enacted in 1975, but until amendments in 1989 it did not require that data be reported at the level of individual applicants. The original intent of the legislation was to simply determine if mortgage applications were being systematically denied due to the location of the property. To some extent this will be correlated with race, but not at a level that is likely to be much use in measuring statistical discrimination. Thus 1990 is the first year of HMDA data on individual applicants available for analysis.

5. Unless otherwise noted all references to "the Boston Fed study" are to Munnell et al. [1996].

6. Note the way in which the bank's determination of the minimum rate required interacts with the assumption of a fixed market rate to determine the approval or denial decision. An alternative model might allow the bank to take on riskier loans, but at a commensurately higher interest rate. This is not the model underlying the Boston Fed study, and justifiably not given it's description of the loan market. But it would change the statistical analysis considerably.

7. There is some slight ambiguity here. The original variable defines six possible actions: loan originated, application approved but not accepted by applicant, application denied, application withdrawn, file closed for incompleteness, or loan purchased by institution. The first two and the last of these are classified here as "approval," and the others as "denial." It is possible that the second action could be interpreted as "denial," since the terms of the loan finally offered might be so unattractive as to amount to a denial of the original loan. A conservative interpretation might classify the first action and the last as approval, and the third as denial, and work with a much smaller data set in which all other applicants were not studied.

8. They can also be readily replicated. An appendix lists the commands used to generate all of our results, using the LMDEP (version 7.0) software documented in Greene [1995]. A machine-readable copy of the command files is available from web site:

http://theweb.badm.sc.edu.proxy.lib.duke.edu/glenn/hmda.htm.

9. Greene [1991, 664-666].

10. One alternative suggested by Greene [1991, 665] is to compute the effect of changes in the binary variable on the probability of denial, using sample values for all other parameters. In our case this makes little difference to the results. The Boston Fed used this method for computing it's marginal effects (see Boston Fed fn. 16 on p. 33 for a discussion).

11. A convenient way to judge the explanatory power of the statistical model is to see how well it predicts the sample behavior. Table

II shows the matrix of prediction errors, using the standard prediction formula whereby an individual is predicted to be denied a mortgage if the logit specification predicts that the probability of him being denied (given his characteristics and the sample coefficient estimates) exceeds or equals 1/2. The "hit rate" can be variously measured; using the main diagonal elements, following the Boston Fed study (see note b to Munnell et al. [1992, Table 5, 27]) we correctly predict 2463 + 152 = 2615 outcomes out of a possible 2931 for a success rate of 89.2%. These equations do, however, underestimate the chance of mortgage denial: the sample probability was 0.145 (= 426 [divided by] 2931) but the predicted probability is only 0.0662 (= 194 [divided by] 2931).

12. Since the sample stratification occurred in relation to an exogenous variable. See Maddala [1983, 171] for further discussion.

13. The first percentage comes from Munnell et al. [1992, Table 1, 6] and the second from Munnell et al. [1992, Table 2, 21]. The public sample used here contains 23.4% blacks and Hispanics.

14. And in Munnell et al. [1992, Appendix B, 49-67].

15. Including the extra variables in this specification reduces the overall sample from 2931 to 1683, since I delete any observation that has missing values for any of the variables included. By expanding the list of explanatory variables one therefore runs the risk of reducing the sample size. It is conceivable that this process could result in certain observations being discarded which generate the difference in estimates, rather than the use of the alternative specification. This is not the case here. If the original Boston Fed specification is re-run on these 1683 observations, rather than the kitchen sink specification, the marginal effect of race on mortgage denial is 5.31% using un-weighted data and 2.81% using weighted data (the standard errors on these effects are, respectively, 1.36% and 1.07%). These estimates match reasonably well with the 6.3% and 3.96% estimates reported in the text using the original data set of 2931 observations.

REFERENCES

Browne, Lynne Elaine. "Boston Fed Study Shows Race Bias." Wall Street Journal, 21 September 1993.

Collet, Maribeth, Glenn W. Harrison, and Thomas F. Rutherford. "Efficient Equity: Removing Salary Discrimination By Meeting Statistical Legal Constraints at Least Cost." Economics Letters, 52, 1996, 81-88.

Greene, William H. Econometric Analysis. New York: Macmillan, 1991.

-----. LIMDEP Version 7.0: User's Manual. Bellport, N.Y.: Econometric Software, Inc., 1995.

Learner, Edward E. Specification Searches. New York: Wiley, 1978.

Liebowitz, Stan. "A Study That Deserves No Credit." Wall Street Journal, 1 September 1993.

Maddala, G. S. Limited-Dependent and Qualitative Variables in Econometrics. New York: Cambridge University Press, 1983.

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

Munnell, Alicia H., Geoffrey M. B., Lynn E. Browne, and James McEneaney. "Mortgage Lending in Boston: Interpreting HMDA Data." American Economic Review, 86(1), March 1996, 25-53.

Harrison: Professor, Department of Economics, University of South, Carolina, Columbia. Phone 1-803-777-4943, Fax 1-803-749-8924. E-mail harrison@darla.badm.sc.edu

---

Please note: Some tables or figures were omitted from this article.