Discrimination and Mortgage Lending in Boston: The Effects of Model Uncertainty

Cullen F. Goenner
Department of Economics
University of North Dakota
PO Box 8369
Grand Forks, ND 58202

E-mail: cullen.goenner@und.edu
Phone: (701) 777-3353
http:/www.business.und.edu/goenner

Abstract:

In 1992 the Federal Reserve Bank of Boston conducted an analysis that examined the effects of race on mortgage lending in the Boston Metropolitan Statistical Area. Collecting data on all possibly relevant information used in the lending process, they find when controlling for a subset of this information that race has a statistically significant effect on the decision to reject a mortgage application. Other researchers, using the same data set, have shown that analysis of alternative subsets of the variables significantly reduces the effects of race. While theory should guide variable selection, there is often no unique theory to explain social science. In such cases, uncertainty in model specification causes one to be uncertain as to the true effects of the variables of interest. This paper accounts for the effects of model uncertainty by using Bayesian model averaging and finds there is little evidence that race has an effect on lending.

(JEL G28, J7)

KEYWORDS: Mortgage lending; Discrimination; Variable selection; Bayesian model averaging.

Discrimination and Mortgage Lending in Boston: The Effects of Model Uncertainty

Homeownership is known to generate many positive economic and social effects on homeowners, while also strengthening neighborhoods and communities. Home equity, as the single largest asset for most Americans, in general has been a good long term investment that provides homeowners with the accumulation of wealth. Purchase of a home using a mortgage also allows homeowners to leverage their investment. While a small fraction of the home's value is used as a down payment, homeowners are able to fully realize any price appreciation in their home. In addition homeownership offers several tax advantages, such as the deduction of mortgage interest. Homeownership though also gives individuals control and responsibility over their own environment, which leads to increased personal satisfaction as well as concern for their community. This results in better citizens as measured by increased electoral participation, lower crime rates, higher graduation rates, and increased family stability.

Purchasing a home for most Americans requires obtaining a mortgage loan. Therefore discrimination by race in the lending process can generate differences in homeownership by race that permeate the economic and social well being of different racial groups as well as their communities. The lending process consists of several stages in which discrimination can take place. Lenders may discriminate in their choice of which neighborhoods to make loans, in their advertising/marketing of products, prescreening of applicants, or the decision to reject a loan application.[1] Discrimination by parties other than the lender, such as real estate appraisers (Schaefer and Ladd, 1981; Schwemm, 1996) and the secondary market (Ross and Yinger, 2002; Van Order, 1996), may also influence access to credit. The focus of this paper is on determining the effects of discrimination in the lender's decision to reject a mortgage loan application. As Ladd (1998) notes, discrimination in the lending process may take place due to reasons other than lender dislike for other races. Pursuit of profit may motivate lenders to use race as a proxy for unobservable characteristics of applicants that influence default. Further white loan officers may have a "cultural affinity" towards white applicants, which causes them to extend greater effort at finding compensating factors that support approval of white loan applications.

Racial differences in homeownership rates are well documented. For the second quarter of 2004 the Census Bureau reported that 76% of whites were homeowners compared to only 50% of blacks, and 48% of Hispanics. Such findings though are not surprising given the racial differences in rejection rates of mortgage loan applications. Data collected as part of the Home Mortgage Disclosure Act (HMDA) indicates for 2003 that white loan applicants were rejected 12% of the time compared to 24% for black and 18% for Hispanic applicants. This disparity in rejection rates continues to exist in the HMDA data when controlling for income. These findings, while troubling, do not necessarily imply that discrimination is the cause of the disparity as other factors relevant to the lending process and correlated with race may be the true cause. For instance minorities may tend to have weaker credit histories, which may explain their higher rejection rate. Discrimination in the lending decision therefore exists when race influences the lending decision after controlling for all the relevant risk factors that influence the profitability of the loan.

In 1992 researchers at the Federal Reserve Bank of Boston conducted an analysis (Munnell et al., 1992) on the effects of race on mortgage lending in Boston. What made this study unique was that Munnell et al. (1996, 43) made the effort to obtain "every variable mentioned as important in numerous conversations with lenders, underwriters, and examiners" to the lending decision. Controlling for a subset of these factors the authors find that race has a statistically significant effect on the decision to reject a mortgage loan, and that this result is robust across several model specifications. The findings were widely discussed among the public, the banking sector, and regulators as to the possible existence of discrimination. Regulators increased exams and the justice department increased scrutiny of mergers and instituted prosecution. The findings though also generated criticism in the popular and academic presses.[2] The primary criticisms were that the results were dictated by the Fed's choice of variables (model specification), data outliers, and simultaneity issues.

In the analysis below the primary focus is on determining the effect of race on the lending decision when accounting for the effects of model specification, though the effect of data outliers is also considered. Uncertainty in model specification occurs when theory is unclear to exactly which variables should be included in the model specification. This may result in researchers who use different variables to come to disparate conclusions over the sign and significance for the coefficients of variables of interest. With respect to the magnitude of the effect of race on lending, Zandi (1993), Day and Liebowitz (1998), and Harrison (1998) each find that using a different subset of the Boston Fed variables greatly reduces the effects of race, which creates uncertainty to the true effect of race. Harrison (1998) notes it is difficult to justify *a priori* why Munnell et al. (1996) exclude many of the variables in their dataset from their models, particularly when each of the variables collected was based on theoretical relevance.

Each of these authors assumes as typical in the literature that the researcher has strong prior information to which combination of variables is the "true" model that generates the data. In this paper this assumption is weakened. We assume that the researcher knows only the list of candidate variables that form the true model, but does not know which combination of these variables form the true model. The candidate variables for consideration are those found in the Munnell et al. (1996) dataset. Using Bayesian model averaging to average over the set of models supported by the data we find that the data does not support the conclusion that race has a significant effect on lending.

Mortgage Lending Decision

The decision to grant a mortgage loan is based on a lender's desire to maximize expected returns, which is influenced by the interest rate and the expected cost from potential default. Lenders though typically do not alter the interest rate charged based on the level of risk, instead they ration credit. The market, which may or may not be competitive, determines the profit maximizing mortgage rate, from which the lender decides to grant mortgages to applicants who are the lowest risk. Stiglitz and Weiss (1981) theoretically motivate this type of credit rationing as due to asymmetric information problems in which the interest rate influences the probability of default. Their argument is that as the interest rate increases, adverse selection increases, generating greater risk. Thus at the market interest rate the demand for credit may be greater than the supply, yet lenders will not charge higher interest rates as expected returns would fall after accounting for higher risk. Williamson (1986, 1987) also ties credit rationing to asymmetric information problems in lending. In the author's model, costly monitoring of loan contracts, rather than adverse selection and moral hazard, are influenced by the interest rate. Higher interest rates increase the probability of default, which increases the cost of monitoring, such that profit maximization need not occur at the interest rate that equates the demand and supply for funds.

With credit rationed, the decision to reject a loan application is negatively related to the expected cost of default, which is the product of the probability and cost of default. One could add to this framework, as Bostic (1996) does, factors that influence prepayment. Determining the probability of default are characteristics of the applicant, property, and the terms of the loan. Factors such as the applicant's income, wealth, occupation, age, and number of dependents influence the economic burden of loan payments, while the age and type of the property along with neighborhood characteristics influence the market value of the collateral and the borrower's decision to default. Loans with higher down payments are less likely to default, as are loans with shorter terms, which build equity more quickly. The cost of default is a function of the value of the collateral (the home) and the terms of the loan. Foreclosure similarly to prepayment results in the need for lenders to reinvest, at potentially lower rates than the original loan terms. The underlying factors that influence prepayment are the same that influence default, though they may have different effects. For example, increasing a household's income may lower the probability of default, but increase the desire for a larger house resulting in prepayment. The probability of a lender rejecting a loan application is thus a function of personal (A) and property characteristics (P), and the terms of the loan (T). $P(R) = f(A, P, T, M)$ Discrimination is said to exist if M (minority) applicants are more likely to be rejected than are whites when controlling for A, P, and T.

Previous Empirical Findings

In 1975 the Home Mortgage Disclosure Act (HMDA) was passed to measure compliance with the Fair Housing Act of 1968. Its stated purpose was to ensure that depository institutions serve the communities where they are located and to determine the distribution by location of public sector investment. The early data indicated that white neighborhoods received five times as many loans as black neighborhoods. This aggregate lending data though failed to account for the supply and demand of credit in these neighborhoods. Schaefer and Ladd (1981) in an early empirical study focused on application data to determine whether discrimination existed in the supply of credit in California and New York. Using data from loan applications for both states the authors examine the lending decision, while controlling for several factors that measure the characteristics of the loan, borrower, property and neighborhood. Their findings indicate, in 22 of the 30 areas examined in California and 6 of the 10 areas in New York, that blacks had statistically significant higher rejection rates than white applicants. In California and New York respectively, whites were 1.54-7.82 and 1.58-3.61 times more likely to be rejected (where statistically significant). These results though were questioned as the authors were unable to control for applicant credit history in both samples and wealth for the California sample. It is well known that failing to include a variable positively correlated with race and negatively related to the lending decision will bias the estimated effects of race as larger than the true effect.

Revisions of the HMDA in 1989 as part of FIRREA demonstrated renewed interest in the issue of race and loan applications. Lenders were now required to collect information on race, gender, income level, and census tract for all mortgage loan applications as well as the disposition of the application, which allowed for the calculation of denial rates by race. Year after year the data show that blacks are more than twice as likely to be rejected as whites. These disparities even remain when controlling for income, causing concern among many, and responses from the lending industry for the need to control for other variables correlated with race and the lending decision.

Towards this end researchers at the Federal Reserve Bank of Boston (Munnell et al. 1992, 1996) examined all applications made by blacks and Hispanics in 1990 and a random sample of whites for the Boston metropolitan statistical area. They asked lenders to provide all the relevant information used in the lending process, which included financial, employment, and property data. In all 38 additional variables, which were noted by lenders, underwriters, and others as theoretically important, were collected for each loan. The data set contains a wealth of information as the researchers have taken great effort to obtain every variable theoretically relevant to the lending decision. Munnell et al's (1996) analysis of a subset of their dataset indicates that after controlling for characteristics of the applicant, property, and loan terms that race has a positive and statistically significant effect on the probability of rejection. Black and Hispanic applicants they find are about 8% more likely to be rejected for mortgage loans than white applicants with the same loan characteristics. The authors use several different combinations of control variables (models) and report the effect of race is consistent across their models.

While the Boston Fed data set was created to reduce the possibility of omitted variable bias, critics argue that the variables selected by Munnell et al. from their dataset influence their findings. The claim is that variables correlated with race, and which determine whether a loan is rejected, are omitted from the analysis causing the estimated effect of race to be biased. Zandi (1993) argues that including four additional variables to the Munnell et al. (1992) model greatly reduces the effect of race.[3] Day and Liebowitz (1998) similarly find in their alternative specification, which also included whether the loan met the lender's credit guidelines and had unverifiable information, that minorities were only 2.8% more likely to be rejected. The effect though remains statistically significant. While researchers have debated over which variables to include, it should be clear as Harrison (1998) notes that the choice of variables in the model is difficult to justify *a priori* as each variable in the Boston Fed data set is by construction relevant. The fear Harrison (1998, 34) adds is whether the model estimated "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." To address this issue Harrison estimates a model that includes the "kitchen sink", which is to say that almost every variable in the data set is included in the model.[4] The finding is that race no longer has a statistically significant effect on the probability of rejection.

Bayesian Model Averaging

Researchers using the Boston Fed data set have a large number of candidate variables to choose from as controls in their models of the mortgage lending decision. Given k candidate variables, there are 2^k different linear models that could be used. Perle, Lynch, and Horner (1993) note the literature provides little guidance as to which measures to include. Researchers are then free to use different subsets of the variables in their models. Existing empirical results, using this data set, indicate that variable selection influences the estimated effects of race. This creates uncertainty as to which model and its results are the true model that generates the data. Rather than accept *a priori* a single model specification as the true model, we examine the entire set of models formed by the different linear combinations of the candidate variables and incorporate our uncertainty into our predictions. Using Bayesian methods, in the form of Bayesian model averaging (BMA), we average the estimated results over the set of models, weighted by the support for each model found in the data, to account for the effects of uncertainty in model specification. For an excellent introduction to BMA see Raftery (1995) and Hoeting et al. (1999), while Brock and Durlauf (2001) and Fernández, Ley and Steel (2001a, 2001b) provide applications in economics.

To begin the researcher must specify the set of models to consider. Here the set of models examined consists of the 2^k different linear combinations of the k candidate variables. The model space for the K models is (M_1, M_2 . . . M_K). Alternative model specifications, such as allowing for interaction terms and different functional form, are not considered. To implement Bayesian model averaging the researcher must specify a prior on the probability that each model is the true model. In the analysis below a uniform prior is used, which assumes that each of the K models is *a priori* equally likely and that $P(M_1) = \ldots P(M_K) = 1/K$. This implies that the prior probability for inclusion of each variable is ½. Fernanez et al. (2001) note this is the standard choice when there is not strong prior information to suggest otherwise. With theory only providing a generalization of which variables to include and lenders with the freedom to weigh factors differently it seems this is a relatively neutral choice without further information. Further, as Raftery (1995) notes, the choice of priors has little influence on the posterior distribution in large samples.

Bayesian methods provide a natural way to estimate the effects of a parameter of interest, such as regression coefficients β, in the presence of model uncertainty. The posterior distribution of β conditioning on the data D is a weighted average of each model's posterior estimates, with the weight being given by the posterior model probabilities $P(M_k/D)$.

$$P(\beta / D) = \sum_{k=1}^{K} P(\beta / M_k, D) P(M_k / D) \tag{1}$$

By Bayes' rule and the law of total probability the posterior model probability is

$$P(M_k / D) = \frac{P(D / M_k) P(M_k)}{\sum_{l=1}^{K} P(D / M_l) P(M_l)} \tag{2}$$

where $P(D/M_k)$ is the likelihood and $P(M_k)$ is the prior probability that model M_k is the true model, which as noted above is assumed to equal 1/K for each model. The posterior model probability then simplifies to:

$$P(M_k / D) = \frac{P(D / M_k)}{\sum_{l=1}^{K} P(D / M_l)} \tag{3}$$

The integrated likelihood is given by

$$P(D / M_k) = \int P(D / \beta_k, M_k) P(\beta_k / M_k) d\beta_k \tag{4}$$

where β_k is a vector of parameters (coefficients and variance), $P(D/\beta_k, M_k)$ is the likelihood and $P(\beta_k/M_k)$ is the prior density of the parameters under model M_k. Raftery (1995) demonstrates using the Laplace method of integrals that the likelihood of model M_k can be approximated as a function, $\exp(-\frac{1}{2} BIC_k)$, of the Bayesian information criterion (BIC) for model k. Schwarz (1978) shows that the BIC is

$$BIC_k = -2 \log(\hat{L}) + d_k \log(N) \tag{5}$$

with $\hat{L}$ equal to the maximized likelihood under model k, d_k is the number of parameters in model k, and N the sample size.

With a large number of candidate variables and thus models, computing equations (1) and (3) requires a great deal of effort given the summations are over the set of K models. Hoeting et al. (1999) describe two means of reducing the computations. The first method, which is used below and by Brock and Durlauf (2001), appeals to Occam's window to discard models that are not supported by the data. The procedure uses the leaps and bounds search algorithm (Furnival and Wilson, 1974) to identify models in the model space with posterior model probabilities significantly worse than that with the highest. Those models below a user specified cutoff are then excluded, and the remaining models are used to average over. Excluding these models has little effect on the posterior estimates given the low weight that each of the models' estimates would receive if included. Raftery (1995) suggests 20 for the cutoff, which is used in the analysis below.[5] The alternative, which is used by Fernandez et al. (2001a, 2001b), is Markov chain Monte Carlo model composition (MC^3). The MC^3 methodology is adapted from Madigan and York (1995) to approximate the posterior distribution of the models based on the models visited by the Markov chain.

When using Bayesian model averaging, the effect of a variable of interest, such as race's effect on the decision to reject a mortgage, can be summarized by its posterior mean, variance, and effect probability. Raftery (1995) reports the posterior mean and variance for β_1 can be approximated by

$$\begin{aligned} E(\beta_1 / D, \beta_1 \neq 0) &\approx \sum_{A1} \hat{\beta}_1(k) P(M_k / D) \\ Var(\beta_1 / D, \beta_1 \neq 0) &\approx \sum_{A1} [Var(k) + \beta_1(k)^2] P(M_k / D) - E(\beta_1 / D, \beta_1 \neq 0)^2 \end{aligned} \tag{6}$$

where $\hat{\beta}_1(k)$ and Var(k) are the maximum likelihood estimates and variance of β_1 under model k, and the summation is over models that include β_1 (set A_1). The posterior effect probability $Pr[\beta_1 \neq 0/D]$ is the posterior probability that β_1 is not equal to zero, which is the sum of the posterior model probabilities for the models that include β_1. The posterior effect probability allows one to evaluate the evidence in favor of a variable having an effect. Raftery (1995, 139) defines the evidence as weak, positive, strong, and very strong based on the breakpoints .5, .75, .95, and .99 on the probability scale, which correspond to Bayes factors of 1, 3, 20, and 150 respectively. For there to be "strong" evidence one must find posterior odds of 20 to 1, similar to the .05 significance level commonly used.

Empirical Analysis

Researchers (Day and Liebowitz, 1998; Harrison 1998; Horne 1997; Munnell et al., 1996; Zandi, 1993) using the Boston Fed data set have found that controlling for different subsets of the variables in the data influences the effect of race on the probability of mortgage rejection. Thus the purpose of this empirical analysis is to account for model uncertainty when estimating the effects of race on the probability that a mortgage loan is rejected. Thirty variables, which are theoretically relevant to the lending decision, are selected from the Boston Fed data set to be candidates for inclusion in the models. Table 1 provides a brief description of the variables. Twenty five of the candidate variables are drawn from Munnell et al.'s (1996) Table 3, which includes the results from five different model specifications using these variables.[6] Added are variables that researchers believe are important and have been omitted, which include the applicant's years of education (Harrison, 1998; Horne, 1997), number of times application reviewed (Harrison, 1998), the amount of liquid assets (Horne, 1997), the presence of unverifiable information (Day and Liebowitz, 1998; Harrison, 1998; Horne, 1997; Zandi, 1993), and whether the applicant met the lender's credit standards (Carr and Megbolugbe, 1993; Day and Liebowitz, 1998; Zandi, 1993; and Horne, 1997).

[Table 1 about here]

Inclusion of the latter two variables in the model's specification has received considerable theoretical and empirical scrutiny. Information on whether each loan application contained information that was unverified and whether each applicant met the lender's credit standards was collected via a survey of lenders a year after the disposition of the loans. With regards to the "credit standards" variable, the Fed researchers (Browne and Tootell, 1995; Tootell, 1996) argue strongly that lenders' responses involved their ex post judgment, which depended on their previous lending decision. The notion they put forth is applicants who are rejected, yet have no credit problems, would be reported by lenders as having failed to meet credit standards. In this case, addition of an endogenous variable would be unwarranted. Day and Liebowitz's (1998) response is that this notion is not clear given that 45% of rejected loan applications met the lender's credit standards. They argue that the variable is important as it may capture differences in lending standards and that independent credit scoring systems are often used to evaluate whether applications meet credit standards. Browne (1996) finds that addition of this variable to their specification reduces the estimated effect of race by one standard deviation, while the effect remains statistically significant at the one percent level.

Unverified information captures whether information on the loan application was verifiable and thus used by the lender in making their decision. Browne (1996) believes inclusion of this variable also could be problematic if lenders respond to the question based on their previous lending decision. The problem, Browne (1996) notes is less severe, as compared to inclusion of the credit standards variable. Carr and Megbolugbe (1993) and Tootell (1996) find that adding this variable to the Boston Fed's model specification reduces the effect of race by a negligible amount and the effect of race remains strongly statistically significant. Thus inclusion of the unverified information variable has not proven "particularly contentious" (Day and Liebowitz, 1998).

In recognition of the debate over inclusion of these two particular variables, we present results from three different sets of models. The first set includes all the linear models formed by the thirty candidate variables (including credit standards and unverified information), the second excludes those models containing the credit standards variable, and the third excludes both the credit standards and unverified information variables from the set of models.

The SPlus program BIC.logit, written by Raftery and Volinsky (1996), is used to implement Bayesian model averaging over the more than one billion logistic regression models formed by the thirty candidate variables. The program reports the model specifications supported by the data, the posterior model probabilities, the posterior mean and standard deviation of the coefficients, as well as the posterior effect probabilities. The results using each of the three model spaces indicate a great deal of uncertainty in the true model's specification. Including the credit standards and unverified information variables resulted in 40 models that were supported by the data, where the model with the highest PMP received only 15% of the total model probability. Table 2 reports the model specifications of the 15 models with the highest PMP from this set of models. The other model sets included 107 and 72 models respectively, with the highest scoring models receiving 6% and 9% of the total model probability.

[Table 2 about here]

Accounting for model uncertainty, we do not find evidence that supports race having a significant effect on mortgage lending in the Boston Fed's sample. For the first set of models the marginal effect of race is computed to be .05%.[7] This implies that minorities are just as likely as whites with the same average characteristics to have their loan applications rejected. The posterior effect probability is 2.1%, which indicates there is evidence that race does not have a statistically significant effect. Results appear in Table 3. The results are quite similar when we exclude the somewhat controversial standards variable. The marginal effect of race is slightly increased to 1.37%, as is the posterior effect probability to 37.1%, which again indicates evidence against race having an effect. This result, unlike previous findings discussed above, demonstrates that including the unverified information variable alone can erase the effects of race when accounting for uncertainty. Excluding both the standards and unverified information variables does increase the magnitude of the marginal effect of race to 4.31% and the posterior effect probability to 87.3%, where the latter number indicates only marginal support for race having an effect.

[Table 3 about here]

A number of variables though did receive strong support, $\Pr[\beta \neq 0/D] > .95$, for having an effect on lending across the three sets of models examined. Debt to income ratio, high LTV ratio, denial of PMI, type of property, unverified information, and credit standards were all found to be important in the first set of models examined. The significance of these variables remains largely the same when the credit standards variable is excluded from the model. The type of property is no longer significant, whereas the number of reviews, medium LTV ratio, consumer credit history, and public record are now significant. Significant variables in the third model set, which excludes credit standards and unverified information, are the same as for the second model set, excluding public record. The sign of these variables, across the three sets of models, is as expected by theory. Several variables did not appear in any of the models averaged over. Of these variables, female is of particular interest. Our results also indicate that gender discrimination does not exist at the lending decision stage of the mortgage process.

To assess the predictive performance of BMA estimates relative to other models, we compare the ability of each to classify as high risk, individuals who in the data are denied loans. For comparison we use the model that receives the highest posterior model probability and the model that includes each of the regressors. To examine out of sample predictions we randomly split the data in half. The first half of the data is used to build the model and obtain estimated coefficients. We then calculate risk scores $\{\exp(x_i^T \hat{\beta})/(1+\exp(x_i^T \hat{\beta})\}$ for each individual in the build data set and define low, medium and high risk groups based on the 1/3 and 2/3 quantiles of risk scores. Using the coefficients from the build data, we calculate risk scores for each individual in the prediction data set and classify each individual to a risk group. Performance is judged by the actual denial rates of individuals assigned to each group, where one prefers that individuals assigned to the high risk group have high denial rates. Table 4 provides results based on the analysis of the set of models that excludes the credit standards variable. One can see that estimates from BMA provide improved out of sample predictions by predicting denial correctly 37.2% of the time as compared to 34.1% for the top model and 35.6% for the full model.

[Table 4 about here]

While the Boston Fed's data set contains a wealth of information, several researchers including Carr and Megbolugbe (1993), Horne (1997), and Day and Liebowitz (1998) have identified several hundred loan applications whose data values appear suspicious. Examples that include five loans accepted despite applicant net worth of a negative million dollars or more. Forty two loans reported as rejected but also sold in the secondary market. Loans that had either unusually low, in some cases negative, or high implied interest rates, relative to the market interest rate. [8] Forty one loans with LTV ratios above 100% and 651 above 80% in which applicants did not apply for primary mortgage insurance. [9] The Boston Fed's response (Tootel, 1996) is largely that the criticism is much ado about nothing. Data outliers they note were double checked with the lender to ensure accuracy and lenders were warned that their data would be turned over to the appropriate regulators. Any "errors" that exist they argue are largely due to decimal errors, which they corrected for in their analysis but not in the released data set, or are errors in variables not used in the analysis, such as whether the loan was sold. The Boston Fed also explains that implied interest rates can be above the market interest rate due to property taxes, which are added to the housing expense, while rates less than market are due to rental property income that offsets expenses. The Boston Fed's arguments though are hardly reassuring, as Day and Liebowitz (1998) find that taxes have a small impact on implied interest rates and that rental income on multi-family properties is unlikely to be the source of low implied rates.

One final check is thus made to determine the effect that outliers may have on the above results. In the analysis below the data set is modified to eliminate observations with implied interest rates that are unreasonable for 1990. Given the mortgage rate averaged around 10% for 1990, we use Carr and Megbolugbe's (1993) criteria to drop loans with implied rates less than 3% or greater than 20%. In addition we drop loans with loan to value ratios greater than 100%, loans reported rejected and sold, and those loans granted in which the lender had net worth less than $-1,000,000. A total of 254 loans are dropped from the analysis. Excluding these observations increases the effects of model uncertainty as the number of models averaged over increases for each of the three sets of models and reduces the effects of race. Results appear in Table 5. The marginal effect of race is found to be .01%, .18%, and 1.21% respectively for the first, second and third model sets. The posterior effect probabilities are also reduced to .6%, 6.8%, and 33.1% respectively, indicating that race does not have even a weak effect across the three sets of models.

[Table 5 about here]

Conclusion

The decision to deny a mortgage loan is complex, as few mortgage applications are perfect. Lenders thus often have the ability to use their own judgment, which may result in different lenders emphasizing different characteristics of the borrower or property in their decision to lend. This implies that there are a large number of theoretically relevant variables that lenders may use. For researchers the large number of candidate variables then implies a rather large space of models for consideration. With thirty potentially relevant regressors there are more than one billion different linear combinations of these variables that researchers may use for their model specification. As is often typical, researchers such as Munnell et al. (1996) report the results from a small number of model specifications, which largely ignores the effects of model uncertainty.

Bayesian model averaging allows researchers a formal treatment of the issue of model uncertainty and can improve predictive performance. With respect to mortgage lending, we find that there is a great deal of uncertainty in the true model specification that generates the data, which is independent of the decision whether to include the credit standards or unverified information variables. Excluding both these variables from the model set, we find 72 models are supported by the data, where the model with the highest posterior model probability explains 9% of the total model probability. Including both of these variables in the set of models examined, results in 40 models that are supported by the data, where the model with the highest posterior model probability explains 15% of the total model probability. Accounting for model uncertainty, the results here indicate it is incorrect to conclude from the Boston Fed data that there is strong evidence to suggest race has an effect on the mortgage lending decision. This conclusion differs from previous findings in that it does not depend on inclusion of the unverified information or credit standards variables or on excluding a large number of observations.

Endnotes

[1] See Goering and Wienk (1996) chapter 1 and Yinger (1996) Chapter 2 for a good review of the potential for discrimination in the lending process.

[2] Ross and Yinger (2003) Chapter 5 provide a thorough discussion.

[3] These variables include whether the applicant's credit history met the lender's guidelines, presence of unverifiable information, cosigner, and the loan amount.

[4] The credit guidelines variable is not included.

[5] Doubling the cutoff did not qualitatively change the results reported below. These estimates are available upon request of the corresponding author.

[6] Munnell et al. (1996) also include census tract and lender dummies in addition to a variable on the rental value of the tract which are not available in the public use data.

[7] The marginal effect of the binary variable race is calculated according to Greene (1997, 878) as $P[Y=1 \mid \bar{x}_*, race=1] - P[Y=1 \mid \bar{x}_*, race=0]$, where $\bar{x}_*$ denotes the means of the other variables.

[8] Given the data set contains the loan amount, the monthly housing expense, and the term of the loan one can calculate an implied interest rate.

[9] Other oddities have been noted in the data set. For a thorough discussion see Tootel (1996) and Yinger (2002).

References

Bostic, Raphael W., 1996. 'The Role of Race in Mortgage Lending: Revisiting the Boston Fed Study.' Unpublished manuscript. Washington, DC, Federal Reserve Board of Governors, December.

Brock, William A. and Steven N. Durlauf, 2001. 'Growth Empirics and Reality.' *The World Bank Economic Review* 15(2): 229-272.

Browne, Lynn Elaine and Geoffrey M.B. Tootell, 1995. 'Mortgage Lending in Boston – A Response to the Critics.' *New England Economic Review*: September/October: 53-78.

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, Theodore E. and S.J. Liebowitz, 1998. 'Mortgage Lending to Minorities: Where's the Bias?' *Economic Inquiry* 36: 3-28.

Fernández, Carmen; Eduardo Ley and Mark F. J. Steel, 2001. 'Model Uncertainty in Cross-Country Growth Regressions.' *Journal of Applied Econometrics* 16(5): 563-576.

Furnival, George. M. and Robert W. Wilson, 1974. 'Regression by Leaps and Bounds.' *Technometrics* 16(4):499-511.

Goering, John and Ron Wienk, 1996. *Mortgage Lending, Racial Discrimination, and Federal Policy*. Washington, DC: The Urban Institute Press.

Greene, William H, 1997. *Econometric Analysis*, 3rd ed. Upper Saddle River, NJ: Prentice Hall.

Harrison, Glenn W., 1998. 'Mortgage Lending in Boston: A Reconsideration of the Evidence.' *Economic Inquiry* 36: 29-38.

Hoeting, Jennifer A.; David Madigan, Adrian E. Raftery and Chris T. Volinsky, 1999. 'Bayesian Model Averaging: A Tutorial.' *Statistical Science* 14(4):382-417.

Horne, David K., 1997. 'Mortgage Lending, Race, and Model Specification.' *Journal of Financial Services Research* 11: 43-68.

LaCour-Little, Micheal, 1999. 'Discrimination in Mortgage Lending: A Critical Review of the Literature.' *Journal of Real Estate Literature* 7: 15-49.

Ladd, Helen F., 1998. 'Evidence on Discrimination in Mortgage Lending.' *Journal of Economic Perspectives* 12(2): 41-62.

Madigan, David and Jeremy York, 1995. 'Bayesian Graphical Models for Discrete Data.' *International Statistical Review* 63(2):215-232.

Munnell, Alicia H; Geoffrey M.B. Tootell, Lynn E. Browne and James McEneaney, 1992. 'Mortgage Lending in Boston: Interpreting HMDA Data.' Federal Reserve Bank of Boston Working Paper 92-07.

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

Perle, Eugene D; Kathryn Lynch and Jeffrey Horner, 1993. 'Model Specification and Local Mortgage Market Behavior.' *Journal of Housing Research* 4(2): 225-243.

Raftery, Adrian E., 1995. 'Bayesian Model Selection in Social Research', in Marsden, P.V., ed., *Sociological Methodology 1995*. Cambridge MA.: Blackwells Publishers (111-195).

Raftery, Adrian E. and Chris T. Volinsky, 1996. Splus function Biclogit, version 2.0. (http://www.research.att.com/~volinsky/bma.html)

Ross, Stephen and John Yinger, 2002. *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement*. Cambridge, Mass.: MIT Press.

Schafer, Rober and Helen Ladd, 1981. *Discrimination in Mortgage Lending*. Cambridge, Mass: MIT Press.

Schwarz, Gideon, 1978. 'Estimating the Dimension of a Model.' *The Annals of Statistics* 6(2):461-464.

Schwemm, Robert G., 1996. 'Housing Discrimination and the Appraisal Industry.' In John Goering and Ron Wienk, eds., *Mortgage Lending, Racial Discrimination, and Federal Policy.* Washington, DC: The Urban Institute Press, pp. 365-397.

Stiglitz, Joseph E. and Andrew Weiss, 1981. 'Credit Rationing in Markets with Imperfect Information.' *The American Economic Review* 71(3): 393-410.

Tootell, Geoffrey M.B., 1996. 'Turning a Critical Eye on the Critics.' In John Goering and Ron Wienk, eds., *Mortgage Lending, Racial Discrimination, and Federal Policy.* Washington, DC: The Urban Institute Press, pp.143-182.

Van Order, Robert, 1996. 'Discrimination and the Secondary Mortgage Market.' In John Goering and Ron Wienk, eds., *Mortgage Lending, Racial Discrimination, and Federal Policy.* Washington, DC: The Urban Institute Press, pp. 335-363.

Williamson, Stephen D., 1986. 'Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing.' *Journal of Monetary Economics* 18: 159-179.

Williamson, Stephen D., 1987. 'Costly Monitoring, Loan Contracts, and Equilibrium Credit Rationing.' *The Quarterly Journal of Economics* 102(1): 135-145.

Yinger, John, 1996. 'Discrimination in Mortgage Lending: A Literature Review.' In John Goering and Ron Wienk, eds., *Mortgage Lending, Racial Discrimination, and Federal Policy.* Washington, DC: The Urban Institute Press, pp.29-74.

Zandi, Mark, 1993. 'Boston Fed's Study Was Deeply Flawed.' *American Banker*, August 19, p. 13.

Table 1: Variable Description

debtinc	Debt to income ratio
concred	1 if no "slow pay" account; 2 if one to two slow pay; 3 if more than two; 4 insufficient credit history; 5 if 60 days past due; 6 if serious delinquencies with 90 days past due
pubrec	1 if any public record of credit problems; 0 otherwise
LTVmed	1 if Loan to value <=.95 and loan to value > .8
LTVhigh	1 if loan to value > .95
pmideny	1 if applicant applied for and was denied PMI; 0 otherwise
nreview	Number of times application was reviewed by lender
unverify	1 if information on the application was unverified; 0 otherwise
selfemp	1 if applicant self employed; 0 otherwise.
housexp	1 if housing expense to income ratio > .3; 0 otherwise
dprop	1 if property 2-4 family home; 0 single family or condominium
race	1 if applicant African American or Hispanic; 0 otherwise
fixrate	1 if fixed rate loan; 0 otherwise
old	1 if applicant age >= MSA median; 0 if applicant age <= median
liqasset	Value of applicants liquid assets (in thousands)
single	1 if the applicant was unmarried; 0 otherwise
school	Years of education
uria	State unemployment rate for applicants industry in 1989
gift	1 if a gift or grant was part of down payment; 0 otherwise
term	Loan term in months
vacancy	1 if tract vacancy > MSA median; 0 otherwise
netw	Value of applicants net worth
mortcred	1 if no late payments; 2 if no payment history; 3 if one or two late payments; 4 if more than two
chval	Change in median value of property in census tract, 1980-1990
boardup	1 if boarded up value > MSA median; 0 otherwise
MHFA	1 if applicant applied under Massachusetts Housing Financing Authority program; 0 otherwise
cosigner	1 if cosigner; 0 otherwise
female	1 if applicant female; 0 otherwise
depend	Number of dependents
standard	1 if applicant met lender’s credit standards; 0 otherwise

Table 2: The 15 Model Specifications with Highest Posterior Model Probability

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
housexp														X	
debtinc	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
pubrec			X	X							X		X		
selfemp	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
LTVmed	X		X		X	X		X	X			X	X	X	
LTVhigh	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
pmideny	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
dprop	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
fixrate	X	X	X	X	X	X			X	X	X	X	X	X	X
old					X							X			
single						X				X	X	X	X		
nreview	X	X	X	X	X	X		X	X	X	X	X	X	X	X
vacancy									X						X
unverify	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
standard	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PMP	14.6	7.8	7.4	6.3	4	4	3.9	3.3	2.7	2.5	2.4	2.4	2.3	2.3	2.2

Table 3: Results of Bayesian model averaging on mortgage lending.

	Model Set Includes Unverify and Credit Standards			Model Set Includes Unverify			Model Set Excludes Unverify and Credit Standards		
Independent Variable	Mean β/D	St Dev β/D	PEP	Mean β/D	St Dev β/D	PEP	Mean β/D	St Dev β/D	PEP
constant	-1.6746	0.521528	100	-5.40612	0.537971	100	-5.2891	0.492957	100
housexp	0.029255	0.121027	6.7	0.359545	0.298542	65.7	0.347887	0.270757	69
debtinc	0.05453	0.009409	100	0.053482	0.010689	100	0.055221	0.009415	100
netw	---	---	0	0.000002	0.000016	1.2	0	0.000007	0.5
concred	---	---	0	0.315957	0.040198	100	0.321394	0.03682	100
mortcred	---	---	0	0.000915	0.016723	0.4	0.029081	0.095523	10.2
pubrec	0.227907	0.347239	34.3	1.439366	0.200565	100	1.204778	0.18709	100
uria	---	---	0	0.003132	0.015888	4.6	0.028094	0.043249	33.7
selfemp	0.773834	0.304852	93.6	0.69851	0.276503	93.5	0.627486	0.277152	90.6
LTVmed	0.311324	0.294806	58.9	0.624703	0.164328	100	0.571496	0.148724	100
LTVhigh	1.724894	0.36175	100	1.844544	0.311798	100	1.631215	0.288161	100
pmideny	4.460226	0.566307	100	4.547448	0.551819	100	4.589863	0.532927	100
dprop	0.77605	0.258956	96.8	0.250916	0.307876	44.7	0.214515	0.280892	41.6
race	0.007177	0.055913	2.1	0.171718	0.245584	37.1	0.45731	0.226879	87.3
boardup	---	---	0	---	---	0	---	---	0
vacancy	0.030846	0.113777	8.3	0.005088	0.042973	1.8	0.001072	0.018022	0.5
chval	---	---	0	0.000004	0.000066	0.4	0.000004	0.00007	0.5
fixrate	0.486805	0.275818	83.4	0.149357	0.228933	34.2	0.020196	0.085039	6.6
MHFA	---	---	0	---	---	0	---	---	0
term	---	---	0	-0.00005	0.000369	2	-0.00002	0.000227	1.1
gift	---	---	0	-0.00919	0.064639	2.6	-0.00334	0.036426	1.2
cosigner	---	---	0	---	---	0	---	---	0
old	0.045375	0.139173	11.4	0.06	0.148073	16.6	0.022684	0.087065	7.8
female	---	---	0	---	---	0	---	---	0
depend	---	---	0	---	---	0	---	---	0
single	0.077226	0.176638	18.8	0.029839	0.105446	9	0.009624	0.055975	3.7
nreview	-0.28227	0.078466	100	-0.2916	0.071948	100	-0.24541	0.064642	100
school	---	---	0	-0.00292	0.013802	5.3	-0.00274	0.01271	5.5
liqasset	0.00001	0.000092	1.5	0.0001	0.000283	13.9	0.000051	0.000198	8
unverify	3.031571	0.275036	100	3.301902	0.250518	100			
standard	-3.56849	0.235589	100						

PEP is the posterior effect probability Pr ($\beta \neq 0/D$).

---These variables were not included in the models that were supported by the data.

Table 4: Evaluating Predictive Performance

	BMA			Top PMP		
Risk Group	Accepted	Denied	% Denial	Accepted	Denied	% Denial
Low	246	7	2.8	369	18	4.6
Medium	576	28	4.6	436	28	6
High	267	158	37.2	284	147	34.1

	Full Model		
Risk Group	Accepted	Denied	% Denial
Low	396	11	2.7
Medium	413	27	6.1
High	280	155	35.6

Table 5: Results of Bayesian model averaging on mortgage lending (Revised sample).

	Model Set Includes Unverify and Credit Standards			Model Set Includes Unverify			Model Set Excludes Unverify and Credit Standards		
Independent Variable	Mean β/D	St Dev β/D	PEP	Mean β/D	St Dev β/D	PEP	Mean β/D	St Dev β/D	PEP
constant	-1.42497	0.587134	100	-5.18837	0.861484	100	-5.44635	0.600368	100
housexp	0.004541	0.047087	1.3	0.195263	0.277974	37.6	0.20372	0.266932	41.7
debtinc	0.047733	0.01006	100	0.053028	0.011077	100	0.054724	0.009842	100
netw	---	---	0	0.000003	0.000022	2.4	0.000001	0.000012	0.7
concred	---	---	0	0.333792	0.042175	100	0.338875	0.039355	100
mortcred	---	---	0	0.002464	0.028727	1	0.033755	0.104762	11.3
pubrec	0.198226	0.335943	29.6	1.519127	0.205214	100	1.309702	0.192857	100
uria	0.002143	0.013947	2.9	0.015862	0.036055	18.9	0.056649	0.052691	59.6
selfemp	0.784665	0.3355	91.5	0.807619	0.243167	98.4	0.720906	0.251803	96.1
LTVmed	0.150859	0.249652	30.8	0.646626	0.186444	99.1	0.581912	0.197937	96.7
LTVhigh	0.016027	0.135055	1.8	0.755992	0.696655	60	0.719676	0.650896	60.8
pmideny	4.962198	0.637621	100	4.866862	0.631996	100	4.873172	0.620951	100
dprop	0.80736	0.291627	95.4	0.151969	0.269436	27.7	0.227779	0.299467	41.3
race	0.001731	0.027238	0.6	0.02706	0.110651	6.8	0.156469	0.244395	33.1
boardup	-0.00544	0.046753	1.8	---	---	0	-0.0002	0.007916	0.1
vacancy	0.001768	0.026244	0.6	---	---	0	---	---	0
chval	---	---	0	---	---	0	---	---	0
fixrate	0.408549	0.31332	70.3	0.084011	0.190103	19.1	0.009998	0.062191	3.2
MHFA	---	---	0	-0.10029	0.315261	11.4	-0.34291	0.529874	34.5
term	---	---	0	-0.00101	0.00184	26.4	-0.00023	0.000893	7.7
gift	---	---	0	-0.0065	0.056282	1.7	-0.00258	0.033157	0.8
cosigner	---	---	0	---	---	0	---	---	0
old	0.125576	0.225722	27.5	0.196836	0.245376	44.4	0.146777	0.213319	36.6
female	---	---	0	-0.00116	0.024896	0.3	---	---	0
depend	---	---	0	---	---	0	---	---	0
single	0.029019	0.112426	7.7	0.014864	0.077086	4.5	0.005576	0.043981	2.1
nreview	-0.19311	0.122945	79.1	-0.28437	0.076378	100	-0.2218	0.079173	96.8
school	---	---	0	-0.0007	0.006609	1.5	-0.00063	0.006231	1.3
liqasset	0.000008	0.000082	1.3	0.000076	0.000248	11	0.00005	0.0002	7.4
unverify	2.950302	0.286738	100	3.198045	0.260027	100			
standard	-3.68086	0.242935	100						

PEP is the posterior effect probability Pr (β≠0/D).
---These variables were not included in the models that were supported by the data.