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

Cullen F. Goenner
Assistant Professor
Department of Economics
University of North Dakota
Grand Forks, ND 58202
e-mail: cullen.goenner@und.nodak.edu
tel: (701) 777-3353
fax: (701) 777-3365

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

Abstract:

In 1992 the Federal Reserve Bank of Boston collected data on all the relevant information used by lenders to evaluate mortgage applications. Controlling for a subset of this information, they find that race has a statistically significant effect on the lending decision. Other researchers, using the same data set, though have shown that analysis of alternative subsets of the variables significantly reduces race's effect. In such cases, uncertainty in model specification causes uncertainty as to the true effects of the variables of interest. This paper uses Bayesian model averaging to account for model uncertainty and finds that the data does not support race having an effect. (JEL G28, J7)

Introduction

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 though 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. The Census Bureau reported in the second quarter of 2004 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, however, 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 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 errors, and simultaneity issues.

Uncertainty in model specification occurs as theory is unclear to exactly which variables should be included in the model. This may result in researchers who use different variables to come to disparate conclusions over the sign and significance for the coefficients of the 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. As Harrison (1998) notes it is difficult to justify *a priori* why Munnell et al. (1996) exclude many of the variables in their data set 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.

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 errors is also considered. We assume that the researcher knows 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) data set. By using Bayesian model averaging to average over the set of models supported by the data we find that the effect of race while positive is reduced when accounting for model uncertainty and that there is evidence that race does not statistically effect the lending decision.

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 probability of 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. As interest rates rise, expected returns rise, as does the probability of default, which increases the cost of monitoring. Similarly, asymmetric information may prevent the interest rate from adjusting to clear the market.

With the mortgage rate determined by the market and credit rationed, a loan's expected return is a function of the probability of default, the cost of default, and the terms of the loan (Yinger, 1996). One could add to this framework, as Bostic (1996) does, factors that influence prepayment. Lenders evaluate the risk associated with a loan based on the personal characteristics of the applicant, those of the property, and the terms of the loan. Applicant characteristics such as income, wealth, occupation, and number of dependents influence the economic burden of loan payments. The age and type of property along with neighborhood characteristics are also important as they influence the value of the collateral and the decision to default. Increasing the down payment and reducing the term of the loan both increase equity, which increase the profitability of a loan. 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 exists if M (minority) applicants are more likely to be rejected than are whites when controlling for A, P, and T.

Lenders need not be racist in order to discriminate. One motivation to discriminate may be to capture the effects of unobservable characteristics that influence the probability of default and are correlated with race. This may result in lenders that utilize different rules of thumb for the information provided by applicants of different races. Schaefer and Ladd (1981) argue this was once the case for women, as lenders treated their income as 50% of a man's income due to perceived differences in future labor participation. Similar behavior may be tied to race, as lenders may perceive that minority applicants have more unstable labor market income or less ability to borrow from family than comparable white applicants. Holding minorities to higher standards based on observed characteristics would result in race having an effect on rejection. Such action while motivated by profitability is illegal.

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 failed to control for applicant credit history in both samples and wealth for the California sample. As is well known, failing to include a variable correlated with race and the lending decision will bias the estimated effects of race.

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 of 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 data set 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 variables omitted by Munnell et al. in their analysis and included in their data set 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 issues 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. As Perle, Lynch, and Horner (1993) note the literature provides little guidance as to which measures to include. Researchers thus may 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 as the true model, in the analysis below we examine the entire set of models formed by the different linear combinations of the candidate variables. By using Bayesian methods (Bayesian model averaging), we average the estimated results over the set of models, weighted by the support in the data for each model, 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 (2001) provide applications in economics.

To begin the researcher must specify the set of models to consider. Each model consists of a different subset of control variables. Here the set of models examined is 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.[5] 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 $P(M_1) = \dots P(M_K) = 1/K$.

This implies that the prior probability for inclusion of each variable is ½. As 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.

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(-½ 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 by Fernandez et al. (2001), 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. The second 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. Raftery (1995) suggests 20 for the cutoff, which is used in the analysis below.

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 measured by its posterior mean, variance, and effect probability. Raftery (1995) reports the posterior mean and variance for β_1 can be approximated by

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

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. Raftery (1995, 146) reports that the evidence in favor of a variable having an effect is weak, positive, strong, and very strong based on the breakpoints .5, .75, .95, and .99 on the probability scale.

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 loan amount (Zandi, 1993), applicant's years of education (Harrison, 1998; Horne, 1997), number of times application reviewed (Harrison, 1998), the amount of liquid assets (Horne, 1997) and the presence of unverifiable information (Carr and Megbolugbe, 1993; Day and Liebowitz, 1998; Harrison, 1998; Horne, 1997; Zandi, 1993).

[Table 1 about here]

Notably absent from this list is the variable that measures whether the applicant met the lender's credit standards. This variable, used by Carr and Megbolugbe (1993), Day and Liebowitz (1998), Zandi (1993), and Horne (1997), has been shown to significantly reduce the magnitude of the effect of race on mortgage rejections. Browne and Tootell (1995) and Tootell (1996) have argued that the credit standards variable does not capture information that was used during the application process, but instead measures information after the loan decision was made. That is lenders answered this question, a year after the disposition of the loan, based on their decision. Day and Liebowitz's (1998) response is that this isn't 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 lenders meet credit standards. While this latter argument is potentially valid, it seems likely that lenders responding to this question would be influenced by their knowledge of their lending decision. Given this doubt the credit standards variable is not included in the list of candidate variables.

Another variable, which has received similar attention and is included, is whether the application contained unverifiable information. The question was designed to capture whether information on the loan application, was verifiable and thus used by lenders in their decision making. This variable was similarly measured of lenders after the lending decision was made and thus as Tootell (1996) notes is potentially endogenous. Both 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 small amount and that the effect of race remains strongly statistically significant. This though is not the case in the presence of model uncertainty as demonstrated below.

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 models supported by the data, their posterior model probabilities, the posterior mean and standard deviation of the coefficients, as well as the posterior effect probabilities. The results provided in Table 2 indicate a great deal of model uncertainty as 106 models are supported by the data and the model with the highest PMP receives only 6.5% of the total posterior model probability.

[Table 2 about here]

Accounting for model uncertainty influences the estimated effects of race on mortgage lending. The posterior mean of race is .173 and the marginal effect is computed to be 1.38%.[7] This implies that minorities are 1.38% more likely than whites with the same average characteristics to have their loan applications rejected. The posterior effect probability is 37%, which as noted above indicates there is no evidence of race having a statistically significant effect. Results appear in Table 3. Other variables that are highly significant, $\Pr[\beta \neq 0/D] > .9$, include debt to income ratio, consumer credit history, public record, self employment, LTV ratios, PMI denial, number of reviews, and whether there was unverified information. The sign of these variables is as expected by theory. Several variables did not appear in any of the 106 models averaged over. Of these variables, female is of particular interest. Our results indicate that gender discrimination does not exist at the lending decision stage of the mortgage process. The posterior means of the coefficients can be used to determine the probability of denial. Using the typical cutoff of .5, where loans with probability of .5 or above are viewed as predicting loan denials, we are able to correctly predict over 90% of application outcomes.

[Table 3 about here]

To assess the predictive performance of BMA relative to other models we compare the ability of each model 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. From Table 4 one can see that BMA provides 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]

To ensure that the results are robust, two modifications were examined. First we consider the effects of including in the list of candidate variables the measure of unverified information. As noted above, while many researchers have included the variable, the researchers (Munnell et al. 1996) who collected the data believe it to be potentially endogenous and therefore argue for its exclusion. After excluding this variable, the results again indicate a large number of models (72) are supported by the data, with the model with the highest posterior model probability receiving only 8.7% of the total probability. Excluding this variable though does increase the posterior mean of race to .458 and the marginal effect to 4.32%. The posterior effect probability of race also increases from 37% to 87%, which according to Raftery's (1995) guide provides positive but not strong evidence for race having an effect on mortgage denials. The significance of the other variables though largely remains the same.

[Insert Table 5 about here]

One final check is made to determine the effect of "problematic" observations in the Boston Fed data set. Carr and Megbolugbe (1993), Horne (1997), and Day and Liebowitz (1998) have all outlined a large number of loan applications whose data values appear suspicious. Examples include loans made with negative implied interest rates, applicant net worth in the negative millions of dollars, loans rejected and then reported sold, and loans granted with loan to value ratios greater than 80% and in which the applicant did not apply for mortgage insurance. For a more thorough discussion see Tootel (1996) and Ross and Yinger (2002).

Without actual access to the original loans it is difficult to determine whether any "errors" are random in nature or are truly problematic. As Browne and Tootell (1995) discuss, unusual observations were double checked with the lender to ensure accuracy and lenders were warned that their data would be turned over to the appropriate regulators. In the analysis below the data set is modified to eliminate observations with implied interest rates that are unreasonable for 1990. The implied interest rate for a mortgage is the interest rate that equates the loan amount to the sum of the present value of the mortgage payments. 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. This differs from the actual interest rate as housing expense includes the mortgage payment in addition to taxes and insurance, and deducts rental income for multifamily homes. 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%. This excludes an additional 178 observations from the data. Dropping these observations has little to no effect on the results reported above (results not shown). With unverified information in the list of controls, 329 models are supported by the data and the posterior mean of race is .137 with a posterior effect probability of 30%. Excluding unverified information results in 113 models averaged over and race having a posterior mean of .48 and effect probability of 89.1.

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 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. By appealing to Occam's window we are able to reduce the number of models under consideration to between 70 and 300. The results indicate a great deal of model uncertainty in mortgage lending as the model with the highest posterior model probability only receives 6% of the total model probability. Inference is based on a weighted average of model estimates, with weight determined by the posterior model probability. Accounting for model uncertainty, the results in this paper indicate that race does not have an effect on the mortgage lending decision. In addition it was demonstrated that Bayesian model averaging can improve out of sample predictive performance.

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, 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, 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, 563-576.

Furnival, George. M. and Robert W. Wilson. (1974). "Regression by Leaps and Bounds." *Technometrics* 16, 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, 382-417.

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

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

Madigan, David and Jeremy York. (1995). "Bayesian Graphical Models for Discrete Data." *International Statistical Review* 63, 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, 25-53.

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

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

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, MA.: MIT Press.

Schafer, Rober and Helen Ladd. (1981). D*iscrimination in Mortgage Lending*. Cambridge, MA: MIT Press.

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

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

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

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

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

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, 135-145.

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

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

Notes

[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 (2002) 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] Researchers Day and Liebowitz (1998) and Horne (1997) both note the potential for non-linear relations in mortgage lending decisions though they both estimate linear model specifications. It is our intention to include model specifications that are typically used in the literature, which are linear.

[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.

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
loanamt	Loan amount (in thousands)

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	X	X	X		X			X	X		X	X		X
debtinc	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
concred	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
pubrec	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
selfemp	X	X	X	X	X	X	X	X	X	X	X	X		X	X
LTVmed	X	X	X	X	X	X	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	
race		X		X	X			X				X			
fixrate				X		X		X	X					X	
old										X					X
nreview	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
unverify	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PMP	6.5	3.8	3.7	3.6	3.5	3.5	3.3	2.7	2.4	2.2	2.2	1.7	1.5	1.5	1.4

Table 3: Results of BMA – Candidate Variables includes Unverify

	Bayesian Model Averaging		
Independent Variable	Mean β/D	St Dev β/D	Pr (β≠0/D) %
constant	-5.4073	0.5385	100
debtinc	0.0535	0.0107	100
concred	0.3158	0.0402	100
pubrec	1.4401	0.2006	100
LTVmed	0.6254	0.1643	100
LTVhigh	1.8454	0.3118	100
pmideny	4.5478	0.5518	100
nreview	-0.2914	0.0720	100
unverify	3.3027	0.2505	100
selfemp	0.6997	0.2763	93.5
housexp	0.3603	0.2988	65.8
dprop	0.252	0.3082	44.9
race	0.1733	0.2463	37.4
fixrate	0.1491	0.2287	34.2
old	0.0590	0.1469	16.4
liqasset	0.0001	0.0003	14
single	0.0303	0.1062	9.1
school	-0.0029	0.0139	5.3
uria	0.0032	0.0160	4.6
gift	-0.0081	0.0609	2.3
term	-0.00005	0.00037	2
vacancy	0.0052	0.0434	1.8
netw	0.000002	0.000016	1.2
mortcred	0.0009	0.0168	0.4
chval	0.000004	0.000067	0.4
boardup*	---	---	0
MHFA*	---	---	0
cosigner*	---	---	0
female*	---	---	0
depend*	---	---	0
loanamt*	---	---	0

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

Table 4: Evaluating Predictive Performance

	BMA			Top PMP			Full Model		
Risk Group	Accepted	Denied	% Denial	Accepted	Denied	% Denial	Accepted	Denied	% Denial
Low	246	7	2.8	369	18	4.6	396	11	2.7
Medium	576	28	4.6	436	28	6	413	27	6.1
High	267	158	37.2	284	147	34.1	280	155	35.6

Table 5: Results of BMA – Candidate Variables excludes Unverify

	Bayesian Model Averaging		
Independent Variable	Mean β/D	St Dev β/D	Pr (β≠0/D) %
constant	-5.2906	0.4933	100
debtinc	0.0552	0.0094	100
concred	0.3213	0.0368	100
pubrec	1.2055	0.1871	100
LTVmed	0.5726	0.1487	100
LTVhigh	1.6324	0.2882	100
pmideny	4.5903	0.5329	100
nreview	-0.2452	0.0643	100
selfemp	0.6283	0.2771	90.6
race	0.4584	0.2267	87.4
housexp	0.3494	0.2708	69.2
dprop	0.2147	0.2810	41.7
uria	0.0282	0.0433	33.8
mortcred	0.02896	0.0953	10.2
liqasset	0.00005	0.0002	8.1
old	0.0222	0.0861	7.7
fixrate	0.0200	0.0847	6.6
school	-0.0028	0.0127	5.5
single	0.0098	0.0565	3.7
gift	-0.0033	0.0364	1.2
term	-0.00002	0.00023	1.1
netw	0.0000004	0.000007	0.5
vacancy	0.0011	0.0182	0.5
chval	0.000005	0.000071	0.5
boardup*	---	---	0
MHFA*	---	---	0
cosigner*	---	---	0
female*	---	---	0
depend*	---	---	0
loanamt*	---	---	0

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