

*Seattle University*  
*Albers School of Business and Economics*

**Equality in Pursuit of the American Dream:**  
The Role of Race in Mortgage Loan Decisions

Afsar Ali, Patrick Hoye, Stefan Rijkaart

ECON 5300

Bridget Hiedemann

March 14, 2018

## 1. Introduction

Owning one's home has long been considered an essential part of the American dream. It's a milestone accomplishment that indicates stability, pride, control, and the opportunity to accumulate wealth and eventually pass it on to the next generation. As a nation of immigrants, we know that many people came to the United States in hopes of achieving such dreams. For most, however, access to credit is crucial to purchasing a home. For this reason, in 1976, the Home Mortgage Disclosure Act (HMDA) was enacted by congress as a means of monitoring minority and low-income access to the mortgage market. The HMDA required that financial institutions provide data involved in mortgage lending decisions to the public. Initially, however, several variables correlated with either race or creditworthiness were not included in these data. In response to this, in 1990, the Federal Reserve Board of Boston requested that financial institutions in the Boston area provide additional information relevant to mortgage lending decisions.

In this study, we examine a sample of mortgage lending decision data from Boston in 1990 to determine whether race is associated with the outcome of a mortgage loan application. Using this data, we seek to answer the question: Do mortgage lending institutions discriminate against minorities? This question is increasingly important as mortgage lending decisions move towards automation via machine learning algorithms. If lending institutions systematically discriminate against minority applicants, the algorithms they employ will learn to do the same.

In order to address our research question, we estimated two separate models, one Probit and one Logit, to determine whether the probability of approval differs among racial groups. Based on the results of our Logit estimate, we found that white applicants faced more than 2 times greater odds of approval than their black or Hispanic counterparts. Similarly, using our Probit model to generate a set of prototypical individuals, we observed that white applicants exhibited higher predicted probabilities of approval than their black or Hispanic counterparts for every combination of demographic characteristics. These results support the notion that racial discrimination exists among the mortgage lending decisions in our sample.

## 2. Econometric Model and Estimation Method

The parameters of our models were estimated using the Maximum Likelihood Estimation (MLE) method. We use this method to determine means and variances based on our available observations, as we are not able to measure data from every member of the population in question. With this information, we estimate Logit and Probit models to determine how each of our variables is associated with the odds and predicted probability of an individual being approved for a mortgage loan.

For both our Logit and Probit models, our dependent variable 'Approved' is binary, with a 1 indicating that the applicant is approved for a mortgage loan, and a 0 indicating that the applicant is not approved. As such, using the values of our independent variables for a particular individual, we can estimate the probability that the applicant will be approved. The independent variables in our model include measures of creditworthiness such as 'Meets guidelines', 'Loan amount / purchase price', and 'Other obligations', as well as personal characteristics such as 'Marital status', 'Black', and 'Hispanic'. We discuss these variables in greater detail in the Data section below.

### 3. Data

For our study, we began with a sample of 1,989 loan applicants collected by the Federal Reserve Board of Boston in 1990, provided by lending institutions in the Boston area. As a relatively small proportion of mortgage loan applications were made by minorities in Boston during this time, the data includes all applications made by black and Hispanic individuals, as well as a random sample of applications made by white individuals. To address our research question, we subset the original data by removing individuals who had null values for marital status, gender and 'meets guidelines'. This eliminated only 20 observations, leaving us with a final sample size of 1,969 applicants.

Variables in the data set include: 'Meets guidelines', 'Loan amount / purchase price', 'Other obligations', 'Marital status', 'Male', 'Black', and 'Hispanic'. 'Meets guidelines' is a categorical variable that indicates whether the applicant meets credit guidelines. 'Loan amount / purchase price' is a percentage that indicates how large of a loan the applicant is requesting relative to the price of their desired home. 'Other obligations' is a percentage that indicates existing financial obligations of the applicant as a percent of their income. 'Married' and 'Male' are categorical variables that indicate whether the applicant is either married or male. Finally, 'Black' and 'Hispanic' are categorical variables that determine whether the applicant identifies as non-Hispanic black or Hispanic, with a 0 in both categories indicating that the applicant is non-Hispanic white.

Table 1 depicts descriptive statistics for the entire sample. We observe that 91% of the sample population met credit guidelines for receiving a mortgage loan, yet only 88% of applicants were approved. Of course, meeting credit guidelines is not the only factor influencing a mortgage loan decision. The average percentage of 'other obligations' for our full sample was 32.39%, ranging from as low as 0% to as high as 95%. In the case of an applicant with other financial obligations that account for 95% of their income, it's understandable that a lending institution would not be keen to approve a mortgage, regardless of whether they meet credit guidelines or not. Average 'loan amount / purchase price' was 77.03% with a median of 80%, indicating that most applicants sought to purchase a home with 20-25% down payment and the remainder as a mortgage. Interestingly, the maximum value for 'loan amount / purchase price' was 257.14%, which is a somewhat suspicious value. We conducted some additional research into this and found that in some cases, an applicant may be able to apply for additional funds to apply towards home improvements or renovations, although in most cases, this would only lead to a loan amount that is marginally higher than the purchase price. Nevertheless, we chose to include these observations, as they constitute a very small portion of the sample. Taking a look at demographic characteristics for our full sample, we see that 66% of applicants were married, 10% of applicants were black, and 5% of applicants were Hispanic.

Table 2 depicts descriptive statistics separated by race/ethnicity. As we know from the descriptive statistics for our full sample, black and Hispanic applicants make up a relatively small proportion of the data set. Thus, our full sample characteristics are greatly swayed by white applicants, who make up the majority. For this reason, it is helpful to examine how characteristics differ among each race. We immediately observe a significant difference in rates of approval, for which white applicants exhibit an average of 91% approved, compared to 67% and 76% for black and Hispanic applicants, respectively. Of course, this does not necessarily lead us to conclude that racial discrimination exists. We observe a similar difference in 'meets credit guidelines', for which 94% of white applicants meet guidelines, compared to 73% and 85% for black and Hispanic applicants, respectively. Average 'loan amount / purchase price' for white applicants was slightly lower than their black or Hispanic counterparts, indicating that, on average, black and Hispanic applicants applied for larger loans proportional to the

price of the home. Interestingly, other obligations as a percent of income were relatively similar among the three groups. White applicants still exhibited the lowest average percentage, at 32.03%, while their black and Hispanic counterparts exhibited averages of 34.90% and 33.47%, respectively. We do not consider this to be an economically significant difference among the three groups. Finally, marital status appears to differ somewhat among the three groups, with Hispanic applicants demonstrating the highest rate of marriage at 71%, compared to 66% and 62% for their white and black counterparts.

**Table 1: Descriptive statistics for full sample**

<b>Variable</b>	<b>Mean</b>	<b>Median</b>	<b>St. Dev</b>	<b>Min</b>	<b>Max</b>
<i>Approved</i>	0.88	1		0	1
<i>Meets credit guidelines</i>	0.91	1		0	1
<i>Loan percentage / purchase price</i>	77.03	80	18.95	2.11	257.14
<i>Other obligations (% of income)</i>	32.39	33	8.28	0	95
<i>Married</i>	0.66	1		0	1
<i>Black</i>	0.10	0		0	1
<i>Hispanic</i>	0.05	0		0	1

**Table 2: Descriptive statistics by race/ethnicity**

<b>Variable</b>	<b>Mean</b>	<b>Median</b>	<b>St. Dev</b>	<b>Min</b>	<b>Max</b>
<b>White</b>					
<i>Approved</i>	0.91	1		0	1
<i>Meets credit guidelines</i>	0.94	1		0	1
<i>Loan amount / purchase price</i>	75.65	79.88	19	2.11	257.14
<i>Other obligations (% of income)</i>	32.03	32.55	8.23	0	95
<i>Married</i>	0.66	1		0	1
<b>Black</b>					
<i>Approved</i>	0.67	1		0	1
<i>Meets credit guidelines</i>	0.73	1		0	1
<i>Loan amount / purchase price</i>	84.06	87.50	17.84	28.99	255.52
<i>Other obligations (% of income)</i>	34.90	35	8.19	5.60	63
<i>Married</i>	0.62	1		0	1
<b>Hispanic</b>					
<i>Approved</i>	0.76	1		0	1
<i>Meets credit guidelines</i>	0.85	1		0	1
<i>Loan amount / purchase price</i>	85.63	89.63	14.5	40.09	162.63
<i>Other obligations (% of income)</i>	33.47	33	8.46	14.60	62
<i>Married</i>	0.71	1		0	1

#### 4. Empirical Results

Tables 3 and 4 depict our Logit and Probit estimates. In both models, we observe that all of our parameter estimates exhibit statistical significance at the 5% level or lower. On this basis, we are able to conclude that our parameter estimates are statistically significant.

Examining the parameter estimates in our Logit model, we see that the coefficients for 'meets credit guidelines' and 'married' are positive, while 'loan amount / purchase price' and 'other obligations' are negative. The coefficient of 3.7214 for 'meets credit guidelines' corresponds to an odds ratio of 41.3217, indicating that an applicant who meets credit guidelines faces roughly 41.32 times greater odds of approval than an applicant who does not, controlling for all other demographic characteristics. Unsurprisingly, this demonstrates that meeting credit guidelines is extremely important in being approved for a mortgage loan. In comparison, the odds ratio for 'married' is 1.5855, indicating that an applicant who is married faces roughly 1.59 times greater odds of being approved than an applicant who is not, controlling for all other demographic characteristics. This makes intuitive sense, as being married might indicate that an applicant is more financially stable. On the other hand, the negative coefficient for 'loan amount / purchase price' corresponds to an odds ratio of 0.9834, indicating that an applicant with 1 percentage point higher loan amount compared to purchase price faces roughly 1.66% lower odds of approval than another similar applicant, controlling for all other demographic characteristics. Similarly, the negative coefficient for 'other obligations' corresponds to an odds ratio of 0.9665, indicating that an applicant with 1 percentage point higher obligations as a percent of income faces roughly 3.35% lower odds of approval than that of another similar applicant, controlling for all other demographic characteristics.

Finally, we examine the parameter estimates for our variables of interest, 'Black' and 'Hispanic', in our Logit model. The coefficients for both variables are negative, corresponding to odds ratios of 0.4442 for 'Black' and 0.4617 for 'Hispanic'. These odds ratios indicate that a white applicant faces roughly 2.25 times greater odds of approval than a black applicant and roughly 2.17 times greater odds of approval than a Hispanic applicant, controlling for all other demographic characteristics. This strongly supports the notion that racial discrimination exists among the mortgage lending decisions in this sample. Interestingly, the odds ratios for 'Black' and 'Hispanic' are relatively similar to each other, indicating that neither race appears to have significantly better odds of approval than the other.

Tables 5 and 6 depict predicted probabilities for prototypical individuals based on our Probit and Logit models. For each model, we examine 12 individuals based on the possible combinations of race, marital status, and whether the applicant meets credit guidelines. For the purpose of comparison, we set 'other obligations' and 'loan amount / purchase price' to their respective sample means. Although the predicted probabilities vary slightly between the two models, the results are quite robust. The largest difference observed between similar prototypical individuals in the two models is 2.18 percentage points, for an applicant who meets guidelines, is single, and is Hispanic.

Examining prototypical individuals based on our Probit model, we observe that the predicted probability of approval for an individual who is white, married, and meets credit guidelines is 89.96%. In comparison, a similar individual who is married and meets credit guidelines but is either black or Hispanic has a predicted probability of approval of either 80.41% or 79.31%, respectively. This gap of roughly 10 percentage points between white applicants and their black or Hispanic counterparts appears to hold constant for each combination of marital status and credit guidelines. At the lowest

predicted probabilities of approval, which are observed among individuals who are single and do not meet credit guidelines, the predicted probability of approval is 13.70% for white applicants, compared to 6.47% and 5.99% for black and Hispanic applicants, respectively. Once again, these results support to the notion that mortgage lending decisions in our sample exhibit racial bias against black and Hispanic applicants.

**Table 3: Logit Estimates**

<i>Variable</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>Odds ratio</i>
<i>Intercept</i>	1.3424 *	0.5667	3.8282
<i>Meets credit guidelines</i>	3.7214 ***	0.2170	41.3217
<i>Loan amount / purchase price</i>	-0.0168 ***	0.0051	0.9834
<i>Other obligations (% of income)</i>	-0.0341 ***	0.0103	0.9665
<i>Married</i>	0.4609 *	0.1810	1.5855
<i>Black</i>	-0.8114 ***	0.2395	0.4442
<i>Hispanic</i>	-0.8973 **	0.3103	0.4077
<i>Number of observations</i>	1,969		
<i>Log-likelihood</i>	-479.73 (df = 7)		

Signif. codes: \*\*\* 0.001; \*\* 0.01; \* 0.05; . 0.10

**Table 4: Probit Estimates**

<i>Variable</i>	<i>Estimate</i>	<i>Standard Error</i>
<i>Intercept</i>	0.5417 .	0.2976
<i>Meets credit guidelines</i>	2.1439 ***	0.1211
<i>Loan amount / purchase price</i>	-0.0084 **	0.0026
<i>Other obligations (% of income)</i>	-0.0164 **	0.0053
<i>Married</i>	0.2289 *	0.0905
<i>Black</i>	-0.4227 ***	0.1266
<i>Hispanic</i>	-0.4617 **	0.1634
<i>Number of observations</i>	1,969	
<i>Log-likelihood</i>	-479.46 (df = 7)	

Signif. codes: \*\*\* 0.001; \*\* 0.01; \* 0.05; . 0.10

**Table 5: Prototypical individuals based on Logit estimates**

	<i>Race/Ethnicity</i>		
	<i>White</i>	<i>Black</i>	<i>Hispanic</i>
<i>Meets guidelines, Married</i>	90.08%	80.14%	78.73%
<i>Meets guidelines, Single</i>	85.14%	71.79%	70.01%
<i>Does not meet guidelines, Married</i>	18.02%	8.89%	8.22%
<i>Does not meet guidelines, Single</i>	12.17%	5.80%	5.35%

Note: 'Other obligations' & 'Loan amount / purchase price' set at sample means. OBRA mean 32.39%; LOANPRC mean 77.03%

**Table 6: Prototypical individuals based on Probit estimates**

	<i>Race/Ethnicity</i>		
	<i>White</i>	<i>Black</i>	<i>Hispanic</i>
<i>Meets guidelines, Married</i>	89.96%	80.41%	79.31%
<i>Meets guidelines, Single</i>	85.32%	73.48%	72.19%
<i>Does not meet guidelines, Married</i>	19.36%	9.90%	9.23%
<i>Does not meet guidelines, Single</i>	13.70%	6.47%	5.99%

*Note: 'Other obligations' & 'Loan amount / purchase price' set at sample means. OBRAT mean 32.39%; LOANPRC mean 77.03%*

## 5. Conclusions

In conclusion, we believe that our results demonstrate an association between race and probability of approval. If we consider that other obligations, loan amount as a percent of house value, and creditworthiness are the three most important factors in a mortgage loan decision, and control for these variables, we find that white applicants exhibit more than 2 times greater odds of approval than their black or Hispanic counterparts. Similarly, prototypical white individuals exhibit predicted probabilities of approval that are 8 to 15 percentage points higher than their black or Hispanic counterparts. For this reason, we are comfortable concluding that some degree of racial discrimination exists in the mortgage loan decisions observed in our data. If this is true, a machine-learning algorithm that attempts to replicate such a decision-making process is at risk of exhibiting racial discrimination as well.

Our study is subject to several limitations. At 1,969 observations, our sample size is relatively small. Of this sample, only 10% and 5% were black and Hispanic, respectively. This means that observations for our groups of interest were limited, and there is greater potential for a few atypical individuals to influence the results. Additionally, although our data contains a few key measures of creditworthiness, we believe that inclusion of additional demographic characteristics would help to refine the model and better observe the association between race and probability of approval. Examples of such characteristics might include applicant age, household/family income, or measures of job stability. Finally, our sample represents a population of Boston-area applicants in 1990. With this in mind, we cannot conclude that our results will hold true for populations outside of this area and time period.

## References

Ladd, Helen F. "Evidence on Discrimination in Mortgage Lending." *Journal of Economic Perspectives* 12, no. 2 (1998): 41-62.

<http://www.csus.edu/indiv/c/chalmersk/econ251fa12/evidenceofdiscriminationinmortgagelending.pdf>

Miller, Jane E., and Yana Van Der Meulen Rodgers. "Economic Importance and Statistical Significance: Guidelines for Communicating Empirical Research ." *Feminist Economics* 14, no. 2 (2008): 117-49.

Smith, Liz. "Top 5 Factors Mortgage Lenders Consider." February 2, 2017.

<https://smartasset.com/mortgage/factors-mortgage-lenders-consider>

## Contributions

Probit/logit models were generated by Patrick Hoye. Descriptive statistics and estimated model tables were created by Stefan Rijkaart. Prototypical individual tables were created by Afsar Ali. Portions of the paper were written and discussed equally by all members, with Patrick Hoye focusing on Data, Stefan Rijkaart focusing on Results, and Afsar Ali focusing on Introduction, Model and Estimation Method, and Conclusion. Stefan Rijkaart proofread and made final adjustments.