Racial Discrimination Among Mortgage Loan Approvals

Free Burton
Thinh Mai
Alan Kuo

Introduction

Loan approval is a crucial part of a capitalistic society if anyone wants to make a life changing decision and go beyond their financial means to purchase a house. Whether you are renovating a house, starting a business, or making an investment in what you believe to be a profitable endeavor, you will most likely need a loan provided through a financial institution. This is especially the case with minorities in many areas, who often have less wealth and income than the average population.

The Home Mortgage Disclosure Act (HMDA) was passed to monitor minority access to the mortgage market. The data includes additional characteristics of each individual and shows that there is a higher denial rate among Black and Hispanic applications than there is for White applications. This is even the case when comparing high income minorities to low income Whites in the 1990s (Munnell, Alicia H., et al. 1996).

It may seem warranted to point out that discrimination is the main cause for the lower approval rate of minorities. However, a further analysis of an applicant's credit history, the ratio of loan amount to purchase price, and other obligations as a percentage of total income may reveal that this discrimination is subject to other factors not included in the 1996 study by Munell, Alicia H., et al. In an attempt to address this question, this paper will analyze the disparity in loan approval rates across races using additional factors captured by banks as a result of the HMDA.

Our model compares Hispanic, Black, and White applicants with identical attributes to consider how their loan approval rates differ. For married individuals who pass credit guidelines, White applicants experience a 95.8% predicted probability of loan approval, whereas Black applicants only have 91.03% and Hispanic applicants only have 90.30%. Even when comparing applicants of the same characteristics, we see a statistically and economically significant gap between White and minority loan approval rates. There are a number of reasons why this difference occurs, such as the blatant possibility of discrimination in financial institution's decisions for loan approval. The following analysis will aim to explore the reasons for these differences.

Econometric Model & Estimation Method

To determine whether race plays a role in the mortgage lending decision, it is necessary to first to account for all factors that might influence the financial institution's decision. Mortgage lenders, like all other businesses, want to maximize their profits. They do so at the time of their lending decision based on all the information given. Their profit is a direct result of the difference between the amount it costs them to lend and their expected return on the loan. Their expected return depends on both the mortgage rate and the probability that a borrower will default. As the mortgage rate is assumed to be consistent with industry standards, that means the probability that a borrower will default plays a large role in lending decisions. It is left up to many characteristics both included and not included by the HMDA dataset. In order to analyze the role that race and ethnicity play in the mortgage lending decision, this study attempts to control for all relevant factors.

In addition, this analysis utilizes both Logit and Probit regression models. In our models, we are estimating the outcome of "Approved", with 0 indicating that the applicant is not approved for a mortgage loan and 1 indicating that the applicant is approved for a loan. We control for a combination of measures that account for the probability of default, loan amount to purchase price, other obligations as a percentage of income, whether the individual meets credit guidelines, their marital status, and their race and ethnicity. These variables will be discussed and elaborated upon further in the data section.

We use the Maximum Likelihood Estimation (MLE) method to estimate the parameters of our models. This method is used to estimate the unknown mean and variance of the population, by finding parameters that maximize the probability of obtaining the observed data. Utilizing these results, we estimate our Logit and Probit models to discover how each of the variables is associated with the odds and predicted probability of an applicant getting approved for a mortgage loan.

Data

This dataset was created as a result of the Home Mortgage Disclosure Act (HMDA), which was enacted to monitor minority and low-income access to the mortgage market. The dataset includes additional variables that are commonly omitted, which allows us to draw stronger conclusions regarding whether minorities are more likely to be denied a mortgage loan than someone who is White. The HMDA data contains variables such as marital status, gender, race, whether mortgage loans were approved, whether credit history meets guidelines, the ratio of loan amount to purchase price of property, and other financial obligations represented as a percentage of total income. The HMDA dataset originally contained 1,989 observations, but we had to subset the data in order to remove incorrect values for marital status, gender, and meeting credit guidelines variables. This left us with a total of 1,969 useful observations that were used in this paper as a representative sample of the Boston area.

The sample data contains information on three major ethnicity groups: White, Black and Hispanic. It includes all applications of Black and Hispanic applicants, and includes a random sample of White applicants. In our analysis, gender, marital status, and race are all represented as dummy-variables. For gender, the female observations represent the reference category. For marital status, the reference category is non-married individuals. For race, the reference category is White. The data focuses on these three ethnicity groups as it represented 94% of the population in the Boston area at the time of data collection.

Table 1 displays descriptive statistics for all race groups and other variables in the total sample. There are some clear indications about the data such as the overwhelming majority of applicants are White and male, and more than half of the applicants are married. In regards to loan approval, on average, 87.6% of applicants in this dataset were approved while 91.3% of applicants passed credit history guidelines.
The fact that White is still the most prevalent ethnic group in this dataset after being sampled speaks to the overwhelming lack of minority loan applicants at the time in Boston. This is clearly represented in our findings, as White applicants represent 84.6% of the total sample. Of the married applicants, 91% are male. Of the non-married applicants, only 62.4% are male.

Table 1: Descriptive Statistics for the entire sample

Variable	Mean (Standard Deviation)
Married	0.659 (0.474)
Meets Credit Guidelines	0.913 (0.282)
Black	0.099 (0.299)
Hispanic	0.055 (0.228)
Male	0.813 (0.390)
Loan Approval	0.876 (0.330)
Other Obligations (% of income)	32.392 (8.279)
Loan Amount / Purchase Price	77.032 (18.949)
Sample Size	1,969

Table 2: Descriptive Statistics for each racial / ethnic group

	White	Black	Hispanic
Variable	Mean (Standard Deviation)		
Married	0.661 (0.474)	0.615 (0.488)	0.713 (0.454)
Meets Credit Guidelines	0.939 (0.240)	0.728 (0.446)	0.852 (0.357)
Male	0.823 (0.382)	0.738 (0.441)	0.796 (0.405)
Loan Approval	0.908 (0.290)	0.672 (0.471)	0.759 (0.430)
Other Obligations (% of income)	32.029 (8.226)	34.898 (8.191)	33.469 (8.458)
Loan Amount / Purchase Price	75.652 (19.009)	84.058 (17.839)	85.629 (14.496)
Sample Size	1,666	195	108

Table 2 represents the descriptive statistics for each of the three ethnic groups in the dataset. While comparing the mean values, there are a few definitive areas that display the difference in how financial institutions perceive different racial groups and their socio-demographic factors. For instance, 93.3% of White applicants from this dataset meet credit guidelines. For Hispanic and Black applicants, that value drops to 85.2% and 72.8% respectively. This difference provides a small glimpse into the different economic and financial backgrounds of these ethnic groups in Boston.

The most interesting statistic however is the percentage of loan approvals. The mean value of White applicants in this dataset that is approved for loans is 90.8%. This value drastically declines for Hispanic and Black applicants at 75.9% and 67.2% respectively. We notice that the difference in the percentage of loan approvals is consistent with the pattern of meeting credit guidelines across race groups. However, it is likely there are factors not accounted for in the data that influence loan approval beyond meeting credit guidelines.

Results

Tables 3 and 4 represent our Logit and Probit parameter estimates respectively. We affirm that each of the estimates are statistically significant at the 5% level or lower across both models.

Table 3: Logit Estimates

Variable	Parameter Estimate (Standard Error)	Odds Ratio
Intercept	1.3424* (0.5667)	3.828
Meets Credit Guidelines	3.7214*** (0.2170)	41.322
Loan Amount / Purchase Price	-0.0168*** (0.0051)	0.983
Other Obligations (% of income)	-0.0341*** (0.0103)	0.966
Married	0.4609* (0.1810)	1.586
Black	-0.8114*** (0.2395)	0.444
Hispanic	-0.8973** (0.3103)	0.408
Sample Size	1,969	
Log-Likelihood	-479.73	

Note: *p < .05. **p < .01. ***p < .001.

In our Logit model, the coefficients for meeting credit guidelines and marital status are positive, whereas the loan amount to purchase price and other obligations as a percentage of income are negative. This is consistent with the signs in the Probit model as well. Controlling for all other demographic characteristics, an applicant who meets credit guidelines will have an odds of approval that is 41.32 times higher than applicants who do not meet credit guidelines. This aligns with the common assumption that one of the first steps for mortgage loan approval is to meet credit guidelines. This is considering that meeting credit guidelines acts as a pre-screening for loan approval, calculating the debt to income ratio, checking credit score, etc.

Moreover, the odds ratio for marital status is 1.586, indicating that an applicant who is married has roughly 1.59 times greater odds of getting approved than an applicant who is single, controlling for all

other variables. Being married obviously does not guarantee an approval for applicants, however, it makes the process to get approval for a mortgage easier. For instance, married couples will apply with both their incomes in a single application, which could help to improve an applicants' debt-to-income ratio. Married applicants might also be considered more financially stable in comparison to single applicants.

In contrast, loan amount to purchase price has a negative coefficient with an odds ratio of 0.983, which indicates that an applicant with a 1 percentage point higher loan amount to purchase price has approximately 1.68% lower odds of getting approved than another applicant, holding everything else constant. This makes intuitive sense, as the higher loan amount can reflect an applicant's income level and how much they can afford for the down payment. Similarly, other obligations has an odds ratio of 0.966 and a negative coefficient of 0.0341, which explains that an applicant with 1 percentage point higher obligations as a percent of income has about 3.5% lower odds of getting an approval than another applicant, keeping everything else constant.

The coefficients for both Black and Hispanic applicants are negative, with odds ratio of 0.444 and 0.408 respectively. These odds ratio indicate that a White applicant has roughly 2.25 times greater odds of approval than a Black applicant and an approximate 2.45 times greater odds of approval than a Hispanic applicant, controlling for all other demographic characteristics.

Table 4: Probit Estimates

Variable	Parameter Estimate (Standard Error)
Intercept	0.5417* (0.2976)
Meets Credit Guidelines	2.1439*** (0.1211)
Loan Amount / Purchase Price	-0.0084** (0.0026)
Other Obligations (% of income)	-0.0164** (0.0053)
Married	0.2289* (0.0905)
Black	-0.4227*** (0.1266)
Hispanic	-0.4617** (0.1634)
Sample Size	1,969
Log-Likelihood	-479.46

Note: *p < .05. **p < .01. ***p < .001.

Table 5: Prototypical Individuals Based On Logit Estimates

	Ethnicity/Race		
	White	Black	Hispanic
Meets guidelines, Married	95.80%	91.03%	90.30%
Meets guidelines, Single	93.51%	86.48%	85.44%
Does not meet guidelines, Married	35.59%	19.70%	18.38%
Does not meet guidelines, Single	25.84%	13.40%	12.44%

Note: Loan Amount / Purchase Price (77.03%) and Other Obligations (32.39%) are set at the sample means. Meeting Credit Guidelines is set by HMDA dataset.

Table 6: Prototypical Individuals Based On Probit Estimates

	Ethnicity/Race		
	White	Black	Hispanic
Meets guidelines, Married	95.88%	90.56%	89.89%
Meets guidelines, Single	93.42%	86.11%	85.23%
Does not meet guidelines, Married	34.20%	20.33%	19.25%
Does not meet guidelines, Single	26.24%	14.49%	13.62%

Note: Loan Amount / Purchase Price (77.03%) and Other Obligations (32.39%) are set at the sample means. Meeting Credit Guidelines is set by HMDA dataset.

Predicted probabilities for prototypical individuals based on our Logit and Probit models are displayed in table 5 and 6 respectively. For each model, we split by race and ethnicity as well as use the possible combination of marital status and meet credit guidelines. We also set loan amount to purchase price to its mean of 77.03% and other obligations as a percentage of income to its mean of 32.39% for comparisons to a typical person.

According to our Logit model, the predicted probability of approval for an individual who is married and meets the guidelines is 95.80% for White applicants compared with a probability of approval of 91.03% for Black and 90.30% for Hispanic applicants. The gap between predicted probabilities for those who meet credit guidelines is largest between White and Hispanic applicants, being roughly 5.5%. However, we observe even larger gaps in loan approval odds between applicants who are White, married, and do not meet credit guidelines. Similar applicants that are Hispanic are 16% less likely to be approved while applicants that are Black are 17% less likely than their White counterparts.

Similarly, we see that if an applicant is White, he or she will have a higher predicted probability for approval regardless of their marital status and credit guidelines requirement. Interestingly, the predicted probabilities of approval for Hispanic and Black are relatively similar to each other, which tells us that neither race appears to have a significant chance of getting approval over the other.

Conclusion

This research paper is set to answer one specific question: Does race discrimination exist among applicants for mortgage loans in the Boston area? Even though White applicants represent 84.6% of the entire dataset, we are still able to find significant patterns to discuss. According to our Logit and Probit models, we find that there exists discrimination among Black and Hispanic applicants for mortgage loan approvals. When examining prototypical individuals, we find that there is a difference in predicted probability of receiving an approved mortgage loan across White, Black and Hispanic applicants, while holding other variables constant. For instance, a married White applicant who passes credit guidelines is 5.5% more likely to receive an approval for a mortgage loan than a married Hispanic applicant who also passes credit guidelines. Black applicants, under the same conditions, experience similar discrimination to Hispanic applicants as they are 4.8% less likely to receive an approved mortgage loan than White applicants.

Given these results, there are still limitations that exist within our analysis. For instance, 84.6% of the data provided in this dataset are White applicants, while Black and Hispanic applicants account for only 15.4%. Thus, the sample provided in the dataset does not accurately represent the population. Furthermore, the data provided is only captured in the Boston area and our conclusions and results cannot be generalized beyond that area. However, it is vitally important to note the real world consequences of such discrimination.

This paper shows that loan approvals have been lined with discrimination as early as the 1990s; this was not that long ago. The question remains, then, have we had time to imprur society and eliminate discrimination in such vital areas as mortgage loan applications? Without access to mortgage loans, minority families will not have the opportunity to move to better neighborhoods that may have stronger education systems and other resources that help the next generation thrive and expand. This would keep them in the cycle of living in lower income neighborhoods and second-tier education systems. In order to eliminate this bias we must consider research such as the analysis done in this paper and examine whether the important decisions we make about other people’s lives are biased.

Citations

Munnell, Alicia H., et al. “Mortgage Lending in Boston: Interpreting HMDA Data.” *The American Economic*

Review, vol. 86, no. 1, 1996, pp. 25–53. *JSTOR*, www.jstor.org/stable/2118254. Accessed 11 Mar. 2020.

Description of criterion Contributions

Free - Free completed the final tables and wrote the majority of the introduction, as well as parts of all the other sections with edits, etc.

Mai - Mai completed the Logit / Probit estimates and wrote the majority of the economic method / model and results section of this paper.

Alan - Alan completed the EDA milestone and wrote the majority of the data and conclusion for this paper.