SEATTLE UNIVERSITY

Mortgage Lending Decisions

by

Ankita Pathak

Nancy Jain

Sanyukta Ghai

ECON 5300

Professor Bridget Hiedemann

March 13th, 2019

CONTENTS

Did the mortgage lending institutions discriminate against minorities in 1990?

Can these results be applied to current issues?

1. Introduction

Housing has widely been acknowledged as an important asset to society. The opportunity to secure a mortgage loan is often the key to an individual's ability to purchase a home. According to Helen Ladd (Ladd, 1998), there has been evidence that many lending institutions in Boston regarded race as a critical factor in making decisions regarding approving mortgage loans. Thus, in many cases, a minority applicant is denied mortgage credit due to aggregate information based on minorities on average, and not only for the individual in question. According to analysis conducted by Bostonplans.org (Mark Melnik, Deputy Director for Research, 2011), number of housing units in Boston have increased tremendously between 1990 and 2010. Also, racial and ethnic composition of Boston has changed a lot since 1990. In 1990, close to 60% of Boston's population was white and 35% was black and Hispanic, whereas in 2010 close to 47% of city's population was white and 40% was black and Hispanic. This means that a large proportion of mortgage loan applicants in Boston are white, black and Hispanic. Hence, in this paper, we want to analyze, whether the mortgage lending institutions have discriminated against minorities in 1990 (based on HMDA data) and whether these results can be applied on the current issues.

Understanding the behavior of mortgage lending institutions towards minority is critical as it will assist us in uncovering any bias in algorithms, that are trained on similar data, which are used to make mortgage lending decisions in recent times. Further if any lending institution discovers evidence of discrimination in their algorithms, they may focus on developing corrective actions which can be used on other loans to monitor and assess their own anti-discrimination efforts. A study conducted by Margery Austin and Felicity Skidmore (Turner & Skidmore, 1999) suggested that lenders should take a harsher stance in foreclosure decisions against minority customers than against whites.

We are using Probit and Logit binary response models to look at the outcome of a mortgage loan application associated with race and ethnicity after controlling for relevant characteristics like other obligations, loan percentage to value of the property and whether the applicant meets guidelines based on the applicant's credit history. According to our findings, predicted probability of approval is more for Non-Hispanic Whites when compared to Non-Hispanic Blacks and Hispanics. Hence, it provides evidence of disparate denials (other things being equal) that establish the presumption that the institutions engaged in discriminatory practices during the mortgage lending process.

2. Econometrics Model and Estimation Methods

We have used binary response models (probit and logit model). Explanatory variables are race and ethnicity, loan to value percentage, other obligations as a percent of income and whether credit history meets the guidelines. Whether credit history meets guidelines can be a direct indicator of the applicant's repayment ability. A timely and appropriate repayment of historical credit can be favorable predictor for future loans and can show that the individual is reliable and has a good financial standing. We have considered other obligations as a percent of total income, since the more the other obligations, it could indicate that the applicant is already financially burdened and may affect their ability of repayment. In case of loan to value percentage, the lower loan to value percent would indicate that the applicant has some amount as down-payment and could show financial prudence or family support in the decision to buy a house. We know that while the financial factors affecting a loan approval are whether credit history meets guidelines, loan to value percentage and other obligations of the applicant as a percent of their income are expected, we want to understand the effect of race and ethnicity in determining the chances of approval for an applicant. We have combined race and ethnicity, and categorized those into non-Hispanic Whites, non-Hispanic blacks and Hispanics. The reference category for race and ethnicity is non-Hispanic whites. Also, "credit history meets guidelines" is the included category and "credit history does not meet guidelines" is the reference category. The predicted probabilities for both probit and logit model cannot fall outside of the 0 - 1 range. Since probit and logit models are non-linear, a parameter estimate does not represent the predicted change in the response probability when the corresponding explanatory variable changes by one unit. The sign of the parameter estimate indicates the direction of the relationship and the p-value indicates the statistical significance of the relationship. The error term follows a standard normal distribution in the probit model and a logistic distribution in the logit model. Both are symmetrically distributed around 0 and independent of the explanatory variables.

Probit Model: The dependent variable is mortgage loan approval which is binary in nature, which means, "success" is "loan approved" and "failure" is "loan denied". Here the response probability is underlying latent continuous random variable, which gives the predicted probability of "mortgage loan approval".

Logit Model: The dependent variable for logit model is log odds of "success" of "mortgage loan approval". If the coefficient of explanatory variable is negative, then odds ratio is less than 1 and if it is positive, then odds ratio is greater than 1.

We have used maximum likelihood method of estimation to obtain parameter estimates.

3. Data

The data used in this paper is mortgage lending data for the state of Boston. This data was collected and compiled by Home Mortgage Disclosure Act in 1990. The Home Mortgage Disclosure Act (HMDA) requires many financial institutions to maintain, report, and publicly disclose loan-level information about mortgages.

3.1 Sample Data

Initial data contains 1989 observations and 8 variables. We have filtered this data based on variables, credit history meets guidelines and loan to value percentage. We have selected only those observations for which guidelines met are either True (Guidelines met = 1) or False

(Guidelines met = 0) and removed arbitrary values like 666. Also, we restrict our sample to loan to value percentage less than or equal to 100 because as per HMDA most of the mortgage lending institutions do not approve loans which are greater than 100%. We have not included marital status and gender in our sample. We have combined race and ethnicity and categorized it into Non-Hispanic Whites, Non-Hispanic Blacks and Hispanics. Variables of our sample data include mortgage loan approval status (approved or denied), credit history meets guidelines, other obligations as a percent of income, loan to value percent, and, race and ethnicity. This leaves us with 1954 observations and 6 variables.

3.2 Sample Descriptive

Below tables provide descriptive statistics for our sample data used in the analysis. In our sample, 8.5% of the loan applicants have credit history that meet guidelines, whereas 91% of the loan applicants have credit history that does not meet guidelines. 85% applicants in our sample are Non-Hispanic Whites. In contrast, only 10% applicants are Non-Hispanic Blacks and 5% of applicants are Hispanics. 12% of the loan applications are denied, whereas 88% of the loan applications are approved. Average approval for Non-Hispanic Whites is 91%, whereas, for Non-Hispanic Blacks and Hispanics, the average approval is 67% and 79% respectively.

Table 1: Descriptive Statistics for Entire sample

Variables	mean	sd	median	min	max
Mortgage Loan Approval	0.88	0.33	1	0.00	1
Credit history meets guidelines	0.91	0.28	1	0.00	1
Other obligations as a percent of total income	32.37	8.22	33	0.00	95
Loan amount as a percent of purchase price	76.08	16.73	80	2.11	100
Non-Hispanic White	0.85	0.36	1	0.00	1
Non-Hispanic Black	0.10	0.30	0	0.00	1
Hispanic	0.05	0.23	0	0.00	1

Note:
Total number of records: 1954

Table 2: Descriptive Statistics for Non-Hispanic White sample

Variables	mean	sd	median	min	max
Mortgage Loan Approval	0.91	0.29	1.00	0.00	1
Credit history meets guidelines	0.94	0.24	1.00	0.00	1
Other obligations as a percent of total income	32.00	8.16	32.50	0.00	95
Loan amount as a percent of purchase price	74.82	17.13	79.86	2.11	100
Non-Hispanic White	1.00	0.00	1.00	1.00	1
Non-Hispanic Black	0.00	0.00	0.00	0.00	0
Hispanic	0.00	0.00	0.00	0.00	0

Note:
Total number of records: 1653

Table 3: Descriptive Statistics for Non-Hispanic Black sample

Variables	mean	sd	median	min	max
Mortgage Loan Approval	0.67	0.47	1.00	0.00	1
Credit history meets guidelines	0.73	0.45	1.00	0.00	1
Other obligations as a percent of total income	35.07	8.10	35.00	5.60	63
Loan amount as a percent of purchase price	82.81	12.59	84.22	28.99	100
Non-Hispanic White	0.00	0.00	0.00	0.00	0
Non-Hispanic Black	1.00	0.00	1.00	1.00	1
Hispanic	0.00	0.00	0.00	0.00	0

Note:
Total number of records: 194

Table 4: Descriptive Statistics for Hispanic sample

Variables	mean	sd	median	min	max
Mortgage Loan Approval	0.79	0.41	1.00	0.00	1
Credit history meets guidelines	0.87	0.34	1.00	0.00	1
Other obligations as a percent of total income	33.31	8.51	33.00	14.60	62
Loan amount as a percent of purchase price	83.48	11.60	88.46	39.39	100
Non-Hispanic White	0.00	0.00	0.00	0.00	0
Non-Hispanic Black	0.00	0.00	0.00	0.00	0
Hispanic	1.00	0.00	1.00	1.00	1

Note:
Total number of records: 107

4. Results

As shown in Table 5 and 6, in our model, the dependent variable is the mortgage loan approval, and the independent variables include credit history meets guidelines, other obligations as a percent of income, loan amount as a percent of purchase price (value), and, race and ethnicity.

Table 5 shows the results of our probit model. It shows the parameter estimates, which we have used to calculate response probability. For both the models in Table 5 and 6, the coefficients for both, other obligations as a percentage of income and racial group Non-Hispanic Blacks are negative and statistically significant at 5% level. By contrast, the coefficient of credit history meets guidelines is positive and statistically significant at 5% level. After controlling for variables like credit history meets guidelines, loan to value percentage and other obligations as a percent of income, on average, predicted probability of loan approval for non-Hispanic Whites is more than that of Non-Hispanic Blacks and Hispanics.

Table 6 shows the results of our logit model. From these, we have used the parameter estimates to calculate the estimated odds ratios. After controlling for other obligations as a percent of income, loan to value percent, and, race and ethnicity, people whose credit history meets guidelines have 42.34 times greater odds of loan approval than people whose credit history does not meet guidelines. On average, for each additional percentage point in other obligations as a percent of income, the odds of loan approval decrease by 3.4%, after controlling for loan to value percent, credit history meets guidelines, and, race and ethnicity. On average, for each additional percentage point in loan to value percent, the odds of loan approval decrease by 1.7%, after controlling for other obligations as a percent of income, credit history meets guidelines, and, race and ethnicity. After controlling for other obligations, loan to value percent and credit history meets guidelines, Non-Hispanic Black loan applicants face 58% lower odds of approval than Non-Hispanic White loan applicants and Hispanic loan applicants face 55% lower odds of approval than Non-Hispanic White loan applicants.

For Table 7 below, we have obtained the predicted probabilities of mortgage loan approval for several prototypical individuals. The table displays predicted probabilities of loan approval, conditional on credit history meets guidelines, and, race and ethnicity, with sample median values of 33% for other obligations as a percent of income and 80% for loan to value percent. As we can see, the predicted probability of loan approval is maximum for Non-Hispanic Whites followed by Hispanics. The predicted probability is the lowest for Non-Hispanic Blacks. This observation is same for both credit history meets guidelines and credit history does not meet guidelines and across both probit and logit models. Also, the gap between predicted probability of Non-Hispanic Whites and Non-Hispanic Blacks and the gap between Non-Hispanic Whites and Hispanics is higher when credit history does not meet guidelines.

Table 7: Predicted Probability of Approval for Prototypic Individuals

		Predicted Probability of Approval	
Credit Guidelines	Race	Probit	Logit
Meets Guidelines	Non-Hispanic White	0.9482	0.9482
	Non-Hispanic Black	0.8808	0.8849
	Hispanic	0.8889	0.8918
Does not meet guidelines	Non-Hispanic White	0.2958	0.3020
	Non-Hispanic Black	0.1621	0.1537
	Hispanic	0.1727	0.1629

Note:
Median Value of other obligations as a percent of income : 33%
Median Value of loan to value percentage : 80%

5. Conclusion

The analysis in this paper implies that there is racial discrimination existing in Boston. We can infer that for two loan applicants with similar financial characteristics, predicted probability of mortgage loan approval is always lower for Non-Hispanic Black and Hispanic applicants in comparison to Non-Hispanic White applicants. This effect remains significant, even after controlling for other obligations as a percent of income, creditworthiness and loan to purchase value percent.

The reason for this disparity could be that, in the past, minorities were less likely to have affluent family members, who could help them in case they got into a financial bind and more likely to be laid off in the event of an economic downturn than whites. Based on an article by Helen Ladd (Ladd, 1998), the motivations of mortgage lending institutions to discriminate against minority borrowers were based on the default rates and incomes of the applicants at that time. After controlling for other obligations, creditworthiness and loan characteristics, these institutions expected minorities to have higher default rates than whites, on average. All other factors held constant, an applicant's income could be used to relate it to the risk of default, with blacks having higher risk of default.

The conclusions above are subject to several limitations. First, although we have not considered marital status and gender for estimation of mortgage loan approval, it points to opportunities for future research. Second, we have categorized race into Non-Hispanic Whites, Non-Hispanic Blacks and Hispanics only. Non-Whites could be separated into more categories to get a better insight about variation in the predicted probabilities of a loan. For example, considering, on average Asians have more income than African American, hence if we separate non-whites further, we can provide better interpretation about the impact of race on predicted probabilities of a mortgage loan approval. Finally, there may be other variables that may affect predicted probability of approval of a loan, for example credit limit, interest rate, installment amount, income, gender, education and age. Including these in the probit and logit model might increase the precision of our estimates as well as eliminate potential omitted variable bias.

Also, based on the report, (Alicia H. Munnell Lynn E. Browne), these effects have not been observed for the same sample over a period of time. For the purpose of this study, our sample is restricted to cross-sectional survey data that is collected and compiled by HMDA for Boston in 1990. Therefore, our results cannot necessarily explain the racial discrimination of mortgage lending institutions across the United States. Also, the number of minority loan applicants is relatively very small compared to Non-Hispanic Whites in our random sample.

To conclude, we believe that our analysis is extremely relevant to current issues. According to Cathy O'Neil in "Can we trust the numbers" (O'Neil, 2018), we are interacting with decision making algorithms all the time. If the historical lending data is biased and is used to train a lending decision algorithm, then that algorithm would continue to have those biases. Such an algorithm could potentially even widen this disparate impact. This means that the use of algorithms in mortgage lending decisions could have unintended consequences — such as shutting the people of color out of the financial system. Because these algorithms can learn on their own, are not strictly governed by rules, and can make their own rules and assumptions, the effects could be proliferated in unexpected ways.

References

Alicia H. Munnell Lynn E. Browne, J. M. (n.d.). Mortgage Lending in Boston: Interpreting HMDA Data. *American Economic Review, 86*, 25-53.

FDIC Law, Regulations, Related Acts. (n.d.). Retrieved from https://www.fdic.gov/regulations/laws/rules/6500-300.html

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

Mark Melnik, Deputy Director for Research. (2011). *Demographic and Socio-economic Trends in Boston: What we've learned from the latest Census data.* Boston Redevelopment Authority, Boston.

O'Neil, C. (2018, January 26). *Can We Trust The Numbers?* Retrieved from Ted Talks: https://www.npr.org/programs/ted-radio-hour/580617765/can-we-trust-the-numbers

Turner, M. A., & Skidmore, F. (1999). *Mortgage Lending Discrimination: A Review of Existing Evidence Mortgage Lending Discrimination: A Review of Existing Evidence.* Boston: The Urban Institute.

Contributions

We have collectively worked through all parts of this assignment in the following manner:

Sanyukta: Responsible for coding of models and predictive probability for prototypic individuals and generating tables. Edited the entire report. Wrote Econometric model section.

Ankita: Interpretation of parameter estimates in Probit and Logit model and Conclusion in the paper. Worked on interpreting the initial descriptive statistics of samples to understand the data.

Nancy: Responsible for writing introduction, econometrics model and estimation method, data and results in the paper and interpretation of logit and probit models.

Appendix

Table 1: Descriptive Statistics for Entire sample

Variables	mean	sd	median	min	max
Mortgage Loan Approval	0.88	0.33	1	0.00	1
Credit history meets guidelines	0.91	0.28	1	0.00	1
Other obligations as a percent of total income	32.37	8.22	33	0.00	95
Loan amount as a percent of purchase price	76.08	16.73	80	2.11	100
Non-Hispanic White	0.85	0.36	1	0.00	1
Non-Hispanic Black	0.10	0.30	0	0.00	1
Hispanic	0.05	0.23	0	0.00	1

Note:
Total number of records: 1954

Table 2: Descriptive Statistics for Non-Hispanic White sample

Variables	mean	sd	median	min	max
Mortgage Loan Approval	0.91	0.29	1.00	0.00	1
Credit history meets guidelines	0.94	0.24	1.00	0.00	1
Other obligations as a percent of total income	32.00	8.16	32.50	0.00	95
Loan amount as a percent of purchase price	74.82	17.13	79.86	2.11	100
Non-Hispanic White	1.00	0.00	1.00	1.00	1
Non-Hispanic Black	0.00	0.00	0.00	0.00	0
Hispanic	0.00	0.00	0.00	0.00	0

Note:
Total number of records: 1653

Table 3: Descriptive Statistics for Non-Hispanic Black sample

Variables	mean	sd	median	min	max
Mortgage Loan Approval	0.67	0.47	1.00	0.00	1
Credit history meets guidelines	0.73	0.45	1.00	0.00	1
Other obligations as a percent of total income	35.07	8.10	35.00	5.60	63
Loan amount as a percent of purchase price	82.81	12.59	84.22	28.99	100
Non-Hispanic White	0.00	0.00	0.00	0.00	0
Non-Hispanic Black	1.00	0.00	1.00	1.00	1
Hispanic	0.00	0.00	0.00	0.00	0

Note:
Total number of records: 194

Table 4: Descriptive Statistics for Hispanic sample

Variables	mean	sd	median	min	max
Mortgage Loan Approval	0.79	0.41	1.00	0.00	1
Credit history meets guidelines	0.87	0.34	1.00	0.00	1
Other obligations as a percent of total income	33.31	8.51	33.00	14.60	62
Loan amount as a percent of purchase price	83.48	11.60	88.46	39.39	100
Non-Hispanic White	0.00	0.00	0.00	0.00	0
Non-Hispanic Black	0.00	0.00	0.00	0.00	0
Hispanic	1.00	0.00	1.00	1.00	1

Note:
Total number of records: 107

Table 5: Probit Model

	Dependent Variable : Mortgage loan approval				
Variables	Estimates	Std. Error	z value	p value	
Intercept	0.640	0.329	1.943	0.052	*
Credit history meets guidelines	2.164	0.122	17.716	0.000	***
Other obligations as a percent of total income	-0.017	0.005	-3.068	0.002	***
Loan amount as a percent of purchase price	-0.008	0.003	-2.390	0.017	**
Non-Hispanic Black	-0.449	0.127	-3.528	0.000	***
Hispanic	-0.407	0.167	-2.439	0.015	**

Note:
*p<0.1; **p<0.05; ***p<0.01
Reference Category for Race is a Non-Hispanic White
Total number of observations: 1954
Log Likelihood: -467.078

Table 6: Logit Model

	Dependent Variable : Mortgage loan approval				
Variables	Estimates	Std. Error	Odds Ratio	p value	
Intercept	1.645	0.668	5.179	0.014	**
Credit history meets guidelines	3.746	0.218	42.344	0.000	***
Other obligations as a percent of total income	-0.035	0.011	0.966	0.001	***
Loan amount as a percent of purchase price	-0.017	0.007	0.983	0.017	**
Non-Hispanic Black	-0.868	0.242	0.420	0.000	***
Hispanic	-0.799	0.321	0.450	0.013	**

Note:
*p<0.1; **p<0.05; ***p<0.01
Reference Category for Race is a Non-Hispanic White
Total number of observations: 1954

Table 7: Predicted Probability of Approval for Prototypic Individuals

		Predicted Probability of Approval	
Credit Guidelines	Race	Probit	Logit
Meets Guidelines	Non-Hispanic White	0.9482	0.9482
	Non-Hispanic Black	0.8808	0.8849
	Hispanic	0.8889	0.8918
Does not meet guidelines	Non-Hispanic White	0.2958	0.3020
	Non-Hispanic Black	0.1621	0.1537
	Hispanic	0.1727	0.1629

Note:
Median Value of other obligations as a percent of income : 33%
Median Value of loan to value percentage : 80%