# Chicago Housing

Annelise Escher

5/29/2022

```
rm(list = ls())
library(tidyverse)
library(stargazer)
library(readxl)
library(knitr)
library(lubridate)
library(dplyr)
library(devtools)
library(sf)
library(janitor)
library(MASS)
knitr::opts_chunk$set(echo = FALSE)
```

## Reading in, cleaning, and combining data

Reading in data

Cleaning main datasets

Adding redlining data and zipcodes

## Data visualizations

Creating dataframe for visualization by zip code

```
## # A tibble: 3 x 7
##   zip   loan_total white_percent_avg black_percent_avg latino_percent_avg
##   <chr>      <dbl>             <dbl>             <dbl>              <dbl>
## 1 60614 7683047000             0.791            0.0368             0.0796
## 2 60647 5446853000             0.530            0.0504             0.346
## 3 60657 5118651000             0.812            0.0279             0.0723
## # ... with 2 more variables: asian_percent_avg <dbl>, majority_race <chr>
```

Recreating WBEZ figure. Map of Chicago zip codes color coded by the dollar amount of mortgage loans made to that area from 2012-2018.

```
## pdf
##   2
```

```
## pdf
##   2
```

```
## # A tibble: 3 x 11
##   census_tract loan_total white_percent_avg black_percent_avg latino_percent_avg
##   <chr>             <dbl>             <dbl>             <dbl>              <dbl>
## 1 17031330100  1747818000             0.482            0.215              0.0611
## 2 17031081800  1010495000             0.826            0.0114             0.0549
## 3 17031833100   835049000             0.606            0.115              0.0892
## # ... with 6 more variables: asian_percent_avg <dbl>, tract_pop <dbl>,
## #   med_family_income <dbl>, avg_loan_amount <dbl>, majority_race <chr>,
## #   id <int>
```

### Creating redlining ranking for 2012-2018 census tracts

```
## [1] 795
```

```
## [1] 742
```

```
##
##   1   2   3   4
##   4  66 382 290
```

```
##
##           1           2           3           4
## 0.005390836 0.088948787 0.514824798 0.390835580
```

Table comparing number of census tracts in each of the four HOLD grades.

```
##
##   1   2   3   4
##   3  61 387 291
```

```
##                  ranking_2010   1   2   3   4
## redline_1930_value
## 1                               0   1   3   0
## 2                               0   6  46  14
## 3                               0  20 222 140
## 4                               3  34 116 137
```

### Regression 1:

```
## Call:
## polr(formula = as.factor(ranking_2010) ~ redline_1930_value,
##     data = for_regression, Hess = TRUE)
##
## Coefficients:
##                     Value Std. Error t value
## redline_1930_value 0.3744     0.1122   3.336
##
## Intercepts:
##     Value   Std. Error t value
## 1|2 -5.3991  1.0540    -5.1223
## 2|3 -1.1960  0.3782    -3.1624
```

```
## 3|4  1.6617  0.3768     4.4103
##
## Residual Deviance: 1351.42
## AIC: 1359.42
```

Table 1:

| | *Dependent variable:* | |
|---|---|---|
| | loans_millions<br>Total home loan amount by Census Tract | |
| | (1) | (2) |
| Continous Redline Grade | $19.532^{***}$ (3.124) | |
| Redline Categories | | $16.785^{***}$ (2.920) |
| Tract Population | $0.014^{***}$ (0.001) | $0.014^{***}$ (0.001) |
| Average Family Income | $-0.012^{***}$ (0.004) | $-0.012^{***}$ (0.004) |
| Average Loan Amount | $0.0002^{***}$ (0.00001) | $0.0002^{***}$ (0.00001) |
| Majority Black | $-14.920^{**}$ (6.126) | $-15.488^{**}$ (6.157) |
| Majority Latino | $-24.126^{***}$ (6.185) | $-24.174^{***}$ (6.211) |
| Majority White | $74.224^{***}$ (5.954) | $73.269^{***}$ (5.967) |
| Observations | 740 | 740 |
| $R^2$ | 0.666 | 0.663 |
| Adjusted $R^2$ | 0.663 | 0.660 |
| Residual Std. Error (df = 732) | 47.880 | 48.069 |
| F Statistic (df = 7; 732) | $208.449^{***}$ | $205.997^{***}$ |
| *Note:* | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Sat, May 28, 2022 - 22:41:53

Table 2:

| | *Dependent variable:* |
|---|---|
| | ranking_2010 |
| redline_1930_value | $0.374^{***}$<br>(0.112) |
| Observations | 740 |
| *Note:* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

### Regression 2: Past category of HOLC grade as predictor for modern recreated distribution

```
##
## Call:
## lm(formula = ranking_2010 ~ value_2 + value_3 + value_4, data = for_regression)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3507 -0.3141 -0.3141  0.6493  0.9000
```

```
##
## Coefficients: (1 not defined because of singularities)
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.35069    0.03646  91.897  < 2e-16 ***
## value_2     -0.25069    0.08246  -3.040  0.00245 **
## value_3     -0.03656    0.04829  -0.757  0.44924
## value_4           NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6188 on 737 degrees of freedom
## Multiple R-squared:  0.01249,    Adjusted R-squared:  0.009806
## F-statistic: 4.659 on 2 and 737 DF,  p-value: 0.009755
```

```
##
## Call:
## lm(formula = loan_total ~ HRS2010 + majority_race, data = for_regression)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -153187631  -22815369   -6976629   17077087  366055369
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           19164792   12713592   1.507    0.132
## HRS2010               14233205    3516583   4.047 5.73e-05 ***
## majority_raceblack   -50092493    6586110  -7.606 8.67e-14 ***
## majority_racelatino -35802063    7202575  -4.971 8.31e-07 ***
## majority_racewhite    93068019    6825262  13.636  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 56890000 on 735 degrees of freedom
## Multiple R-squared:  0.5264, Adjusted R-squared:  0.5238
## F-statistic: 204.2 on 4 and 735 DF,  p-value: < 2.2e-16
```

https://stats.oarc.ucla.edu/r/dae/ordinal-logistic-regression/