

# CIFP\_Rough\_Draft

Joey Herrera

4/23/2021

## Causal Inference Final Project Rough Draft

### Introduction (Points of emphasis that will be fleshed out)

- There is a growing literature in redlining that insinuates cultural districts catalyze gentrification in their region instead of creating sustainable economic development.
- This paper seeks to use a weighted difference in difference estimation to find the effect that cultural districts have on average housing prices in metropolitan regions in Texas.
- This paper will add to the literature on redlining and gentrification by
- The inspiration of this paper comes from the creation of the Master Plan zoning policy in Austin in 1928 that segregated all families of color and low-income families into a six square mile radius. In 2013, the city of Austin decided to name this area a cultural district and mitigate the displacement of families of color through gentrification.

## Data Methodology

In order to find the effects that cultural districts have on the average housing prices in major Texas metropolitan regions, I crafted a data sample from IPUMS, which contains micro level data acquired through the Census and the American Community Survey. From this data extract I took key indicators of housing price including census tract, individual income, household income, race, gender, and others from 2008-2018. During this ten year period I also extracted data from Zillow's Typical Housing Price Value Index, which provides typical housing prices on different metropolitan regions in Texas. Finally, I created a crosswalk to map unique metropolitan identification numbers from the IPUMS data to the region identification numbers Zillow uses in their API. The resulting data frame consists of approximately 200 observations which accounts for seventeen different regions from 2008 to 2018.

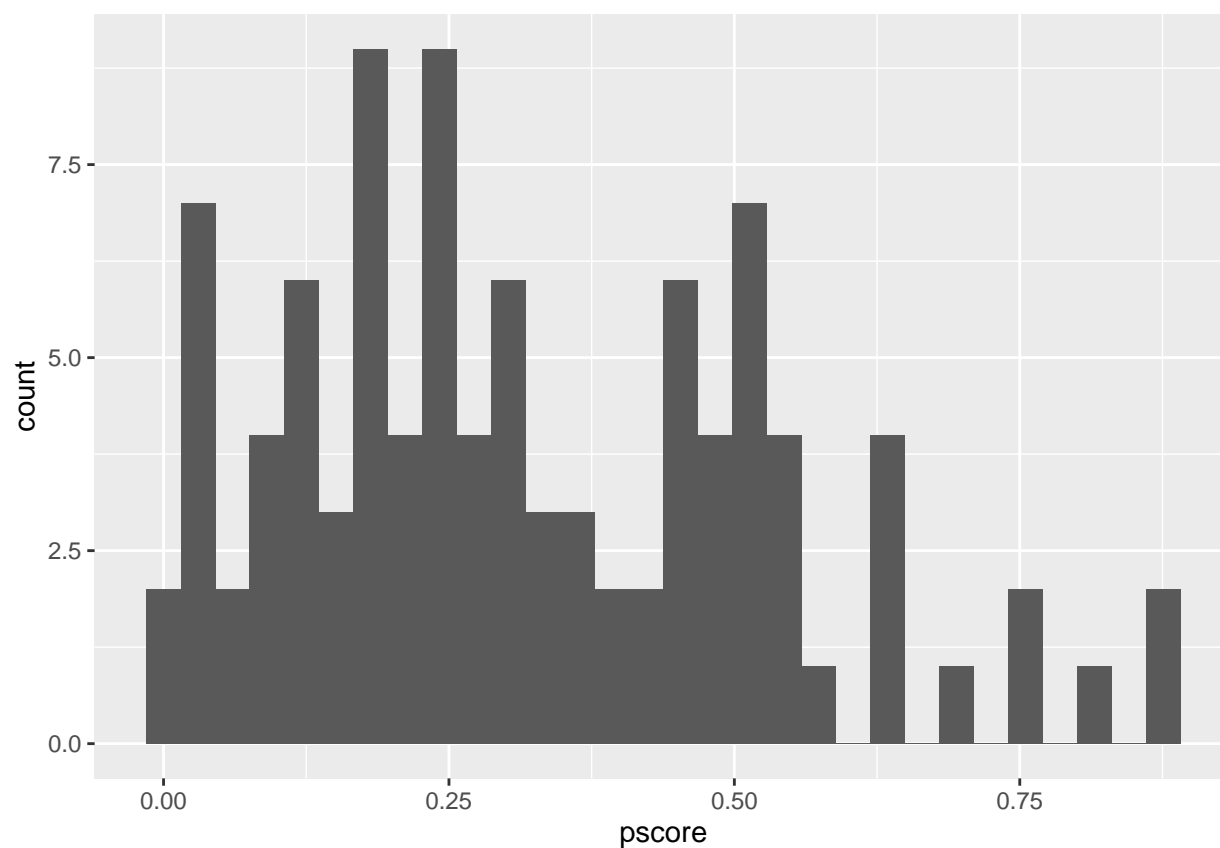
## Empirical Methodology

To begin my empirical methodology, I created a dummy variable called `treat`. When `treat` equals 1, or the year is greater than or equal to 2013, the observation will be included in the treatment group. If `treat` equals zero, then the observation will be in the control group.

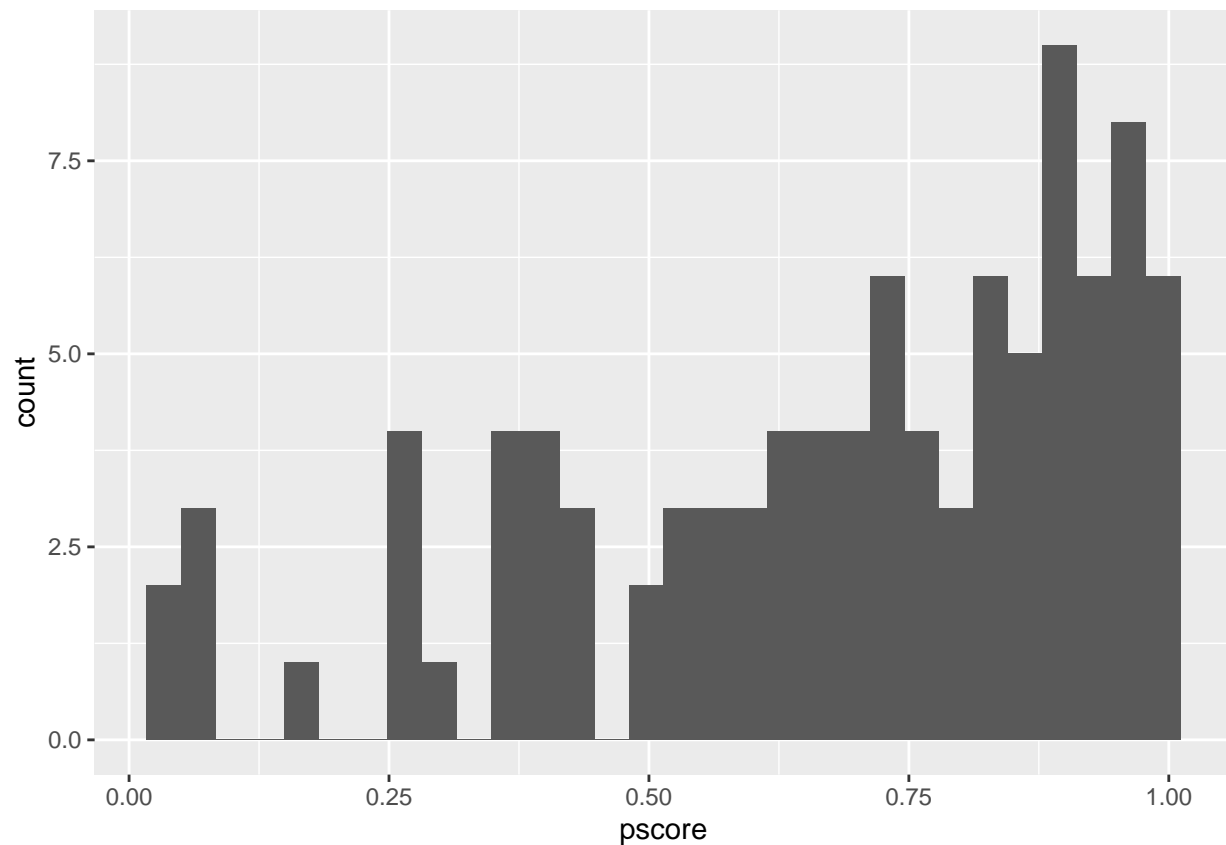
The weights in the weighted difference in difference regression are created by fitting a logit model with the dependent variable `treat` on to the other covariates, which creates propensity scores. The propensity score for the treatment group is approximately 0.68, while the propensity score for the control group is about 0.32.

The histograms below depict the propensity scores frequency for the treatment and control groups. Both histograms agree with the propensity scores that the logit regression estimates, which are an average of the propensity scores from the treatment and control groups. The histogram illustrating individual propensity scores from the control group show a higher density in the 0 to 0.2 range. This is unlike the results from the treatment histogram, which depicts a higher density of individual propensity scores above the 0.5 mark.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

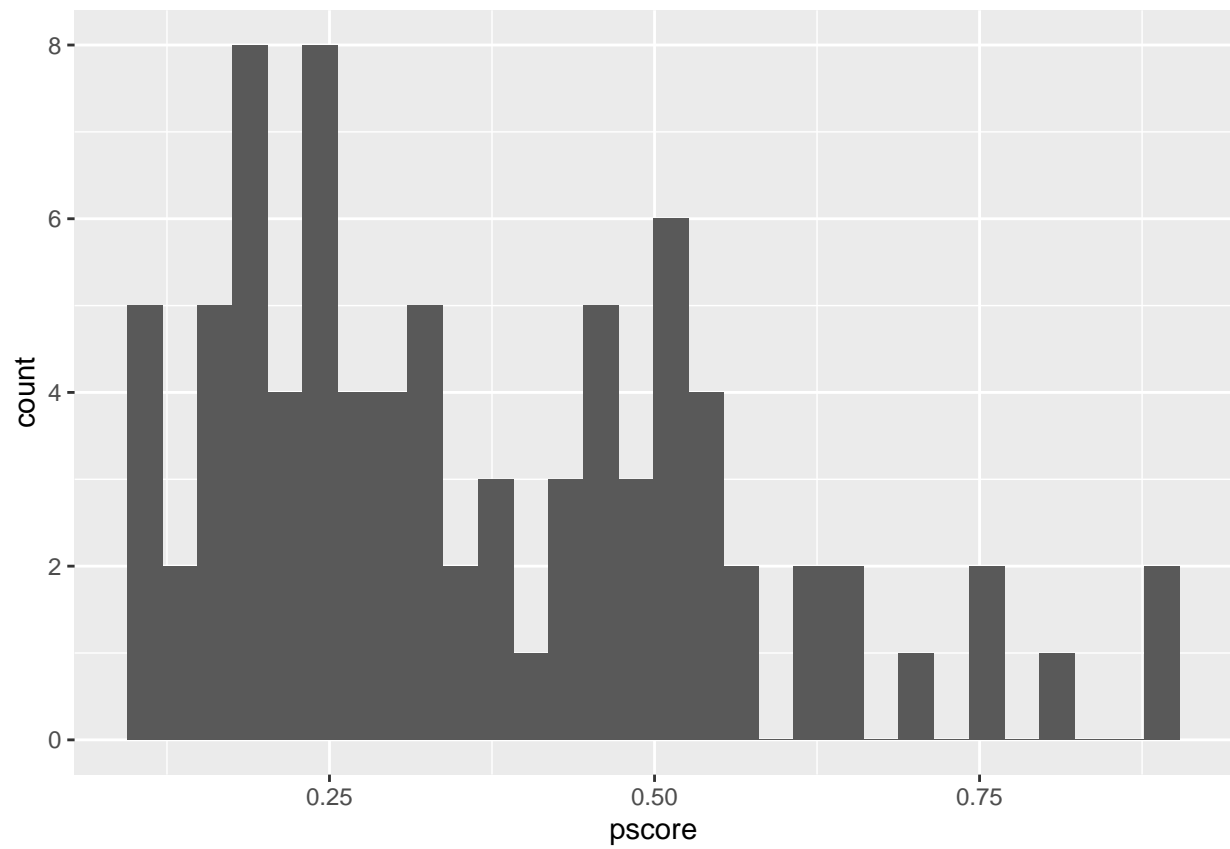


```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

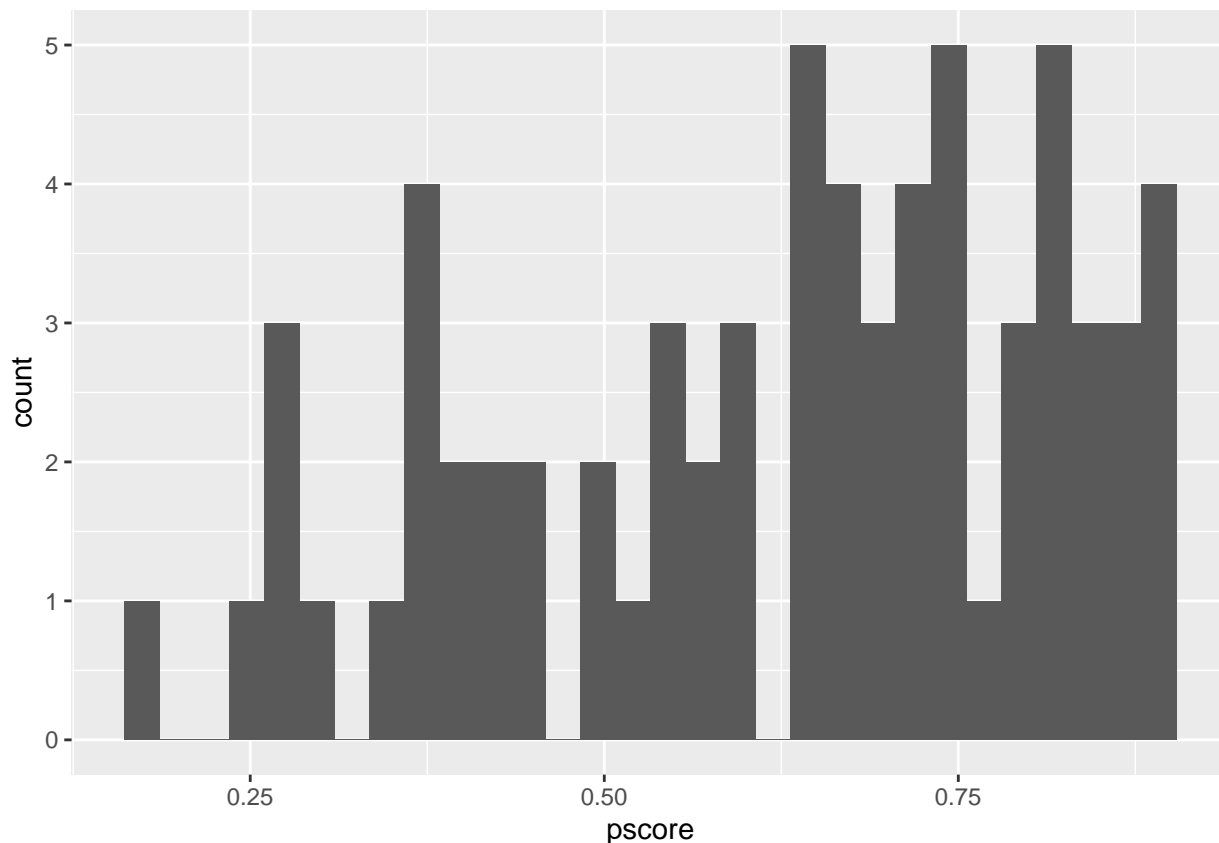


In order to get more accurate average propensity scores for the treatment and control groups I trimmed the data by filtering out individual propensity scores below 0.1 and above 0.9. The resulting average propensity score for the control group grew to 0.37 from 0.32. The trimmed average propensity score for the treatment decreased to about 0.63 from 0.67. The trimmed individual propensity scores are reflected in the two histograms below, which are identical to the previous two histograms except they do not include the observations that were recently omitted.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



The first difference in difference regression this paper performs is a simple two-by-two difference in difference estimator. The purpose of estimating this regression is to compare the simple version to the propensity weighted version. The resulting average treatment effect is 18315.67 dollars. Thus, the simple difference in difference estimator determines that the effect of cultural districts on average housing prices in metropolitan regions in Texas is approximately 18,315 dollars. This differs from the weighted difference in difference estimator estimats that the effect cultural districts have on average housing prices in metropolitan regions in Texas is 281,852 dollars.

```
## # A tibble: 68 x 1
##   avg_price
##   <dbl>
## 1  117451
## 2  225836.
## 3   99440.
## 4  154648.
## 5  120573
## 6  152633.
```

```

## 7 148129
## 8 115518.
## 9 114756.
## 10 98307.
## # ... with 58 more rows

## # A tibble: 80 x 1
##   avg_price
##   <dbl>
## 1 112085.
## 2 89973.
## 3 155221.
## 4 119403.
## 5 149063.
## 6 146576.
## 7 135158.
## 8 140314.
## 9 117276.
## 10 215996.
## # ... with 70 more rows

## [1] NA

```