Food Deserts and Health Outcomes

Hannah Jones

4/23/2021

Using Differences-in-difference Design to Determine the Effect of Food Access on Obesity in the US

In order to study the effect of food insecurity on impacted populations, I used grocery store distance data from 2010 and 2015, and health data from within the time period of change (2013) and after (2017). I then examine health outcomes for these treated to groups to see if increased distance from groceries causes poor health outcomes compared to those groups with no change in status.

The primary goal of this paper is to measure the causal effect between access to food and health outcomes. The identifying assumption underlying this differences-in-difference strategy is the changes in health outcomes for counties with small changes in access to food provide a good counterfactual for the changes in health outcomes that would have been observed for counties with larger changes in access if their access to food had changed similarly.

To assess this assumption, I focus on the distance to the closest grocery store, plotting health outcomes over time for groups sorted according to their changes in distance to the closest store between 2010 and 2015. The data contains flags to indicate which census tracts are within 0.5 miles or 1 mile of a grocery store for urban areas, and withing 10 or 20 miles for rural areas. The movement of these flags between time periods will serve indicate grocery store openings if distance reduces, and grocery store closings if the distance increases for any tract. For the outcome variable, we will look at health data collected in 2013 and 2017. The markers of interest that are generally associated with low food access are obesity, diabetes, high blood pressure and high cholesterol. The data from the CDC surveys indicate the prevalence of obesity, diabetes, high blood pressure, and high cholesterol as a percentage of the population.

Using the distance indicators, I designate the treatment group as any rural or urban population that become a food desert between the two observation years. This is to say that in the first set of data from 2010, the census tract observed has a grocery store within 1 mile if it is urban, and within 20 miles if it is rural, but then in the second set of data, the census tract no longer has a grocery store within the appropriate distance. This "treatment" indicates the closure of grocery stores since the distance to the closest store has now increased outside of the 1-mile (urban)/ 20-mile (rural) threshold.

To run a differences-in-difference model, we use the emergence of a food desert (grocery store closures) as the treatment, and we can use the prevalence obesity as the outcome. We will compare the prevalence of obesity in areas that retained their proximity to grocery stores to these areas that emerged as food deserts between 2010 and 2015. We assume parallel trends in this analysis. This assumption requires that nothing unobserved changes within a census tract that would also determine obesity prevalence.

Analysis:

To analyze the impact of food desert status on health, I use a weighted dif-in-dif technique. I predicted a propensity score for each Census Tract based on population, urban status, income status, and a variety of health factors. I chose these covariates to capture the propentsity of already low-income and low-health populations to lose grocery store access.

```
## # A tibble: 2 x 4
```

```
##   treat   mean    max      min
##   <dbl>  <dbl>  <dbl>    <dbl>
## 1     0 0.0636  0.995 0.000354
## 2     1 0.184   0.993  0.00243
```

Above are the mean propensity scores for the treated and untreated census tracts. I dropped the pscores over 0.9 and under 0.1 to remove extreme cases.

Propensity Scores for Untreated

Propensity Scores for Treated
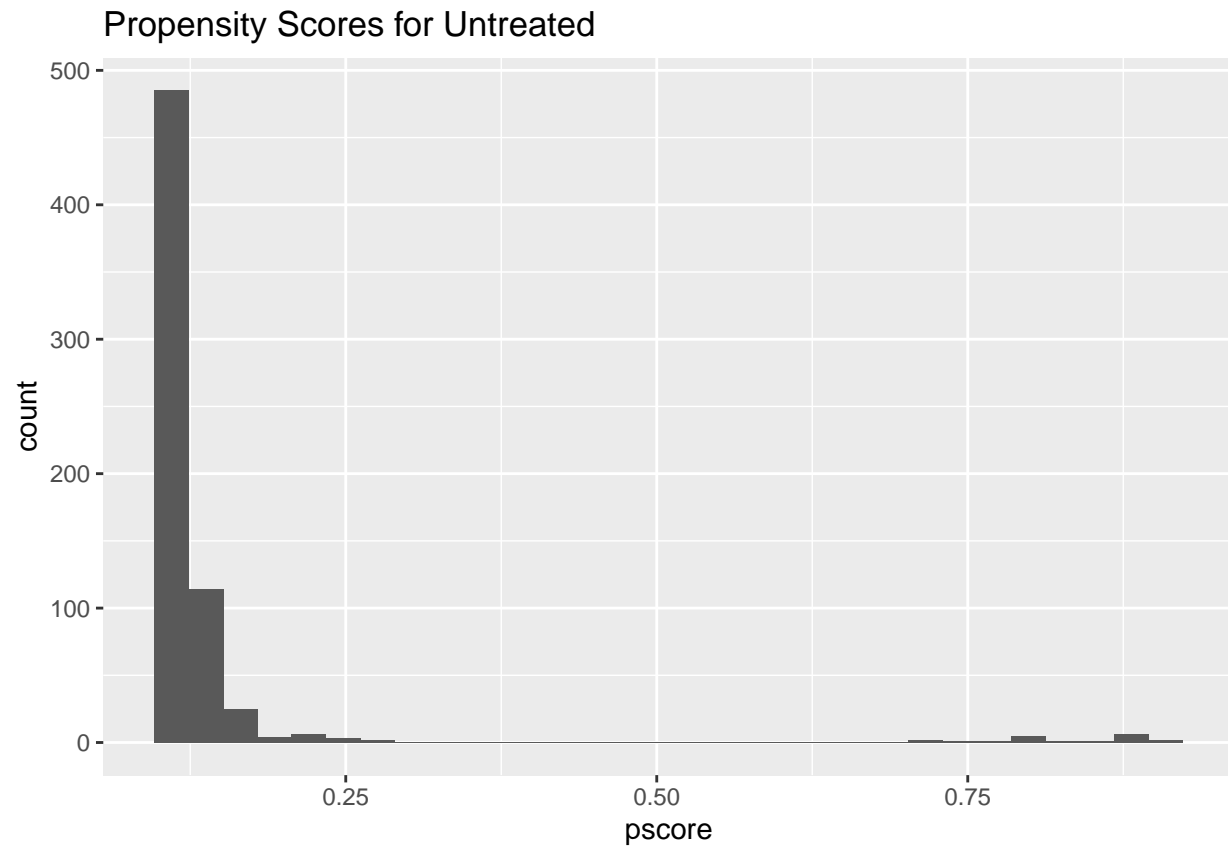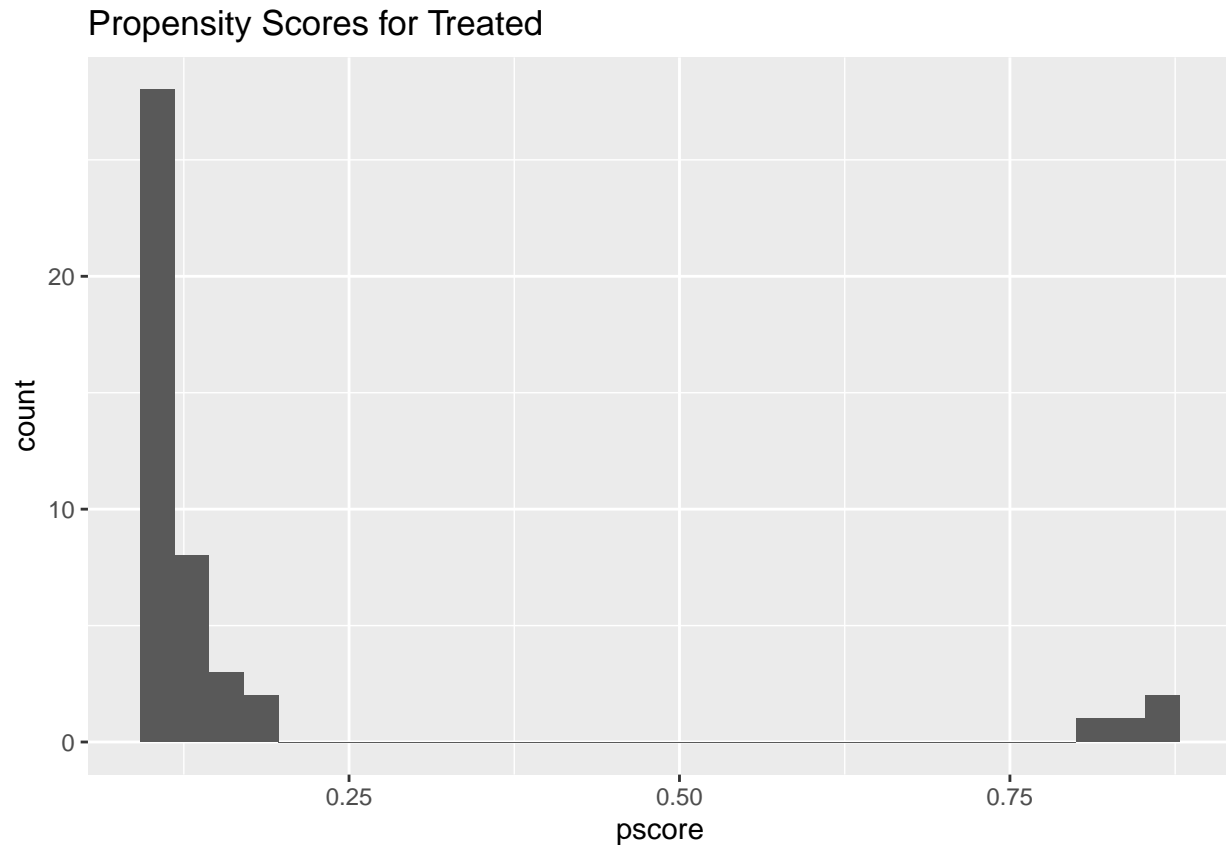

```
## # A tibble: 2 x 4
##    treat logit_mean logit_max logit_min
##    <dbl>      <dbl>     <dbl>     <dbl>
## 1      0      0.138     0.899     0.100
## 2      1      0.182     0.861     0.100
```

The histograms above show the distribution of propensity scores for the treated and untreated Census Tracts. The summary shows that the mean propensity score is slightly higher for the treatment group. We will use these propensity scores to weight the treatment effect. First, we can look at the average treatment effect of grocery store closures between 2010 and 2015 on obesity prevalence between 2013 and 2017. As shown below, the average treatment effect of grocery store closure is and increase in obesity prevalence of ~4.4%.

```
## [1] 4.408994
```

Now I will look at the weighted difference using the propensity scores, using both non-normalized and normalized weights.

```
## [1] -16.16335
```

Above is the non-weighted difference in obesity prevalence. This suggests that grocery store closures result in a 16.6% decrease in the prevalence of obesity. This is extremely high and counter-intuitive, but once we normalize the weights, we find a difference of 4%. This suggests that the difference in the prevalence of obesity in a census tract that has experienced grocery store closures is about a 4% increase over a tract that has not experienced closures.

```
## [1] 4.196173
```