

Week 7 Homework

DATA 210

Data: Don't forget that all data we provide you for this class can only be used for class purposes.

1. In many states, almost all states convicted felons are banned from voting while they serve their prison, parole, or probation sentence. In some of those states, they are able to have their voting rights restored. Even after voting rights are restored, former felons continue to register to vote and vote at extremely low rates. In this question, we will look at an experiment (co-authored by UPenn's Dr. Marc Meredith) that sought to understand what caused these low rates. The researchers were specifically interested in how much of this low rate could be explained by felons' not knowing that their voting rights had been or could be restored after serving their sentence.

For these exercises, you will use the dataset 'felons.RData' to replicate some of the findings in the article "Can Incarcerated Felons be (Re)integrated into the Political System? Results from a Field Experiment."¹

- a. Imagine you wanted to understand the effect that felony convictions and incarceration have on voter registration and turnout (after the sentence has been served). Somebody suggests that you compare the voter registration rates and turnout rates of former felons to people who never have served time for a felony. Discuss why this research design could or could not help you to get a good estimate of the causal effect that you are interested in.
- b. Read the **Experimental Design** section of the "Can Incarcerated Felons be (Re)integrated into the Political System?" (pages 915 through 917 in *gerber, et al 2015.pdf*).
 - i. What are the causal effect(s) that the authors are interested in studying?
 - ii. Describe the treatment and control conditions in the experiment.
 - iii. Describe the randomization strategy that the authors used.
- c. Now we're going to analyze the results from the experiment. Begin by removing the 161 people in the dataset who returned to prison before the experiment was conducted. Then create a new variable called 'treatment_collapsed' which tells us whether each observation in the data was in the control group (FALSE) or a treatment group (TRUE).
- d. The first thing you should always do before analyzing the results of an experiment is assess whether you have balance in your treatment and control groups. In a well-balanced experiment, no pre-treatment covariates (i.e. the variables that existed before you ran the experiment) would predict whether or not somebody ended up in the treatment or control group. For the following questions, use the `treatment_collapsed` variable.
 - i. Use 4 t-tests to assess whether the felons' age, number of days served in prison, time since their release from prison, or 2008 vote turnout is a statistically significant predictor of treatment. To do the t-tests, you'll want to write code that looks like this: `t.test(felons$age ~ felons$treatment_collapsed)`. Create a well-formatted table the present the average values for each of these variables in the treatment and control groups, as well as the the p-value associated with the difference between those averages. You can pull out these values from the output of the `t.test()` object using the `$` operator. Is there significant imbalance for any of those four variables?
 - ii. Use linear regression to assess whether the type of crime predicts whether somebody ended up in the treatment or control group. Were any crimes strong predictors of the treatment?
 - iii. Use linear regression to assess balance for all the variables (age, days in prison, time since release, 2008 turnout, crime type) simultaneously. When you do this, do you find imbalance

¹<https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12166>

for any of the pre-treatment covariates?

- e. Did the experiment have an effect on whether or not ex-felons registered to vote? Did it impact their turnout in 2012? If so, how much did the treatment increase or decrease the probability that they registered or turned out? You can use linear regression and the ‘treatment_collapsed’ variable to answer this question.
- f. Use linear regression to estimate these two treatment effects again. This time, control for the five pre-treatment covariates (the ones you checked for balance in part C in your regression. What effect did the treatment have on registration and voting?