

Replication # 2

1. How does recreational marijuana legalization affect crime.

Several states have legalized marijuana for recreational use. The goal in this study, is to estimate whether recreational marijuana legalization has increased violent crime, based on data from the supplemental homicides reports (SHR) (shr.csv file),

Start with micro level data from SHR. In the dataset, I would you like to keep only murders involving 3 or fewer victims to avoid mass shootings or other crimes (terrorism events like 9/11). So keep only the murders with additional_victim_count of 3 or fewer. Aggregate up to the state by year level. Then merge with population data from the state_pop.csv file. Create a variable murder_pop which is murders/(pop/100000). The average murder rate in the US is around 5 recently. Inspect the time series for each state. You'll find you should drop 3 state_fips regions due to under-reporting/missing data (you find which 3).

The following states legalized recreational marijuana recently.

Colorado – 2014
Washington -2014
Oregon – 2015
Alaska – 2015
Nevada - 2017

- a. Create a dummy variable that indicates whether recreational variable is available within a state. Now run a regression with $\log(\text{murder_pop})$ on the left. On the right hand side, control for state fixed effects, year fixed effects, and the recreational marijuana dummy. What is it's coefficient, and upper and lower bounds. Does marijuana contribute to reefer madness as suggested by Alex Berenson in his recent book?
- b. Now I want you to collapse your data to 3 regions. Colorado, Washington, and the rest of the country. Create a connected line plot with murder on the y-axis and year on the x-axis. Create a vertical line marking 2014. Broadly happens to average murder rates in Colorado and Washington, the first legalizers? What happens in the rest of the country? Do the trends in the Colorado and Washington match those in the rest of the country?
- c. Now use the state by year data again. Drop California, Oregon, Nevada, and Alaska (they legalized early, but not enough for post treatment data). Use a synthetic control approach to estimate the effect of marijuana legalization for both Colorado, and Washington. Match on the lagged murder rate per 100,000 for each year 2000, 2001, 2002, etc. through 2013. How does the murder rate in Colorado and Washington compare to its synthetic control average for 2014-2017?

- d. Use randomization inference to conduct hypothesis testing. To do so, reestimate the synthetic control approach for each state not treated excluding the two treatment states (and the other later adopting states). Estimate the mean squared error (MSE) for the placebo states, and for the treated states, both both and after recreational marijuana legalization. Now generate a ratio of $\text{postMSE}/\text{preMSE}$. How does CO and Washington compare to the distribution. This is your empirical p-value. Can you reject the H_0 that marijuana legalization did not increase crime? Why or why not?

2. Quantile Regressions

Reopen the dataset you already downloaded for replication 1, `LATE_BETTER_THAN_NEVER.csv`. Now make a histogram as you did before for the distribution of individual treatment effects. Make sure you put vertical lines at the 5th percentile of treatment effects, the median, and the 95 percentile of treatment effects.

Now I want you to estimate quantile treatment effects. You can either do this on your own, or use the QTE package. <https://cran.r-project.org/web/packages/qte/vignettes/R-QTEsWrapper.pdf>

Estimate QTE's for every percentile from the 1 percentile to the 99th percentile. Do QTE's allow you to recover the distribution of individual treatment effects?