

The Effects of Marijuana Legalization on Crime and Other Drug Usage in Seattle

Benjamin Gaiser and Jeremy Russell

16 Dezember 2016

Contents

Introduction (400 words?)	1
Literature Review and Policy Relevance (600 words?)	1
Data Gathering and Methodology (1000 words?)	1
References	3

Introduction (400 words?)

Lorem ipsum dolor sit amet, oportere pertinacia repudiandae ne mea, eam ne meis eleifend. Ferri splendide tincidunt cum ad, vis mandamus posidonium cu, eos no vitae sententiae honestatis. Tantas inciderint sed eu, aequae sonet efficiantur no sed. Porro debet pertinax per at, ea mea unum vide.

Literature Review and Policy Relevance (600 words?)

Lorem ipsum dolor sit amet, oportere pertinacia repudiandae ne mea, eam ne meis eleifend. Ferri splendide tincidunt cum ad, vis mandamus posidonium cu, eos no vitae sententiae honestatis. Tantas inciderint sed eu, aequae sonet efficiantur no sed. Porro debet pertinax per at, ea mea unum vide.

Data Gathering and Methodology (1000 words?)

Data Gathering

Our project will be using three main data sets: i) cannabis data on Washington State, ii) crime data on Seattle, WA, and iii) socio-economic data on Seattle, WA. As the policy field of cannabis legalization is a rather nascent field of study, there has been limited availability of reliable data (previous data sources were clouted by the illegality of possession and consumption and thus approximated dark figures).

Our project has taken advantage of one of the provisions within the ‘Washington Initiative 502’, the initiative that brought about the legalization of cannabis in Washington State, that established disclosure of retailers, processors and producers as well as some general economic data online. The data has been web-scraped using the rvest package (Wickham 2016). As there have been difficulties due to the 502data site being scripted in Java, we additionally needed the ‘phantomjs’ file to use the RSelenium package (Harrison 2016) to make the web-scraping possible.

As the “exe”-file does not work on Macs and because scraping is time-consuming, we will provide an unaltered “txt”-file called “ProdProc” for all producers and processors and one called “Retailer” for retailers. This way we allow for as much reproducibility as possible. The data lists the names and locations of producers/processors in the state and of retailers including YTD sales and tax revenue. For the later analysis, we will constrain the data set to Seattle. The second set of data concerns data on crime. We scraped the data for Seattle with the jsonlite package (Ooms, Temple Lang, and Hilal 2016) and formatted it into R-readable

format with the gdata package (Warnes et al. 2015). The initial data set is rather large (>80 MB). We have transformed it into a zipped-file called “CrimeSeattle.gz”. For the analysis, we cleaned the data so to only contain information about the longitude/latitude, the type of crime and the time it has occurred. For the types of crime we have aggregated individual types into seven broad categories: i) alcohol-related crime (includes Liquor offenses and DUI offenses), ii) burglary/theft, iii) marijuana-related crime, iv) narcotics-related crime, v) other-drug related crimes, vi) property crime, and vii) violent crime (includes armed robberies, assaults, drive by shootings, strong arm robbery and homicides).

The third set of data comes from the U.S. Census data and provides us with socio-economic factors for each district. One limitation is that we, however, only have this data annually and only for 2014. The variables included for this analysis are the share of poverty for each district, the average age level, the prevalence of minority races, as well as the level of education. While these are classic socio-economic factors, the lack of variation means that they will only be useful for the regular OLS regression as well as for the propensity score matching, which means that this set of data will be of limited value overall.¹

Methodology

¹should find quote for stereotypical soes

References

- Harrison, John. 2016. *RSelenium: R Bindings for Selenium WebDriver*. <https://CRAN.R-project.org/package=RSelenium>.
- Ooms, Jeroen, Duncan Temple Lang, and Lloyd Hilaiel. 2016. *Jsonlite: A Robust, High Performance JSON Parser and Generator for R*. <https://CRAN.R-project.org/package=jsonlite>.
- Warnes, Gregory R., Ben Bolker, Gregor Gorjanc, Gabor Grothendieck, Ales Korosec, Thomas Lumley, Don MacQueen, Arni Magnusson, Jim Rogers, and others. 2015. *Gdata: Various R Programming Tools for Data Manipulation*. <https://CRAN.R-project.org/package=gdata>.
- Wickham, Hadley. 2016. *Rvest: Easily Harvest (Scrape) Web Pages*. <https://CRAN.R-project.org/package=rvest>.