Estimate the Impact of Opioid Control Policies

(Report for Nick)

White Team

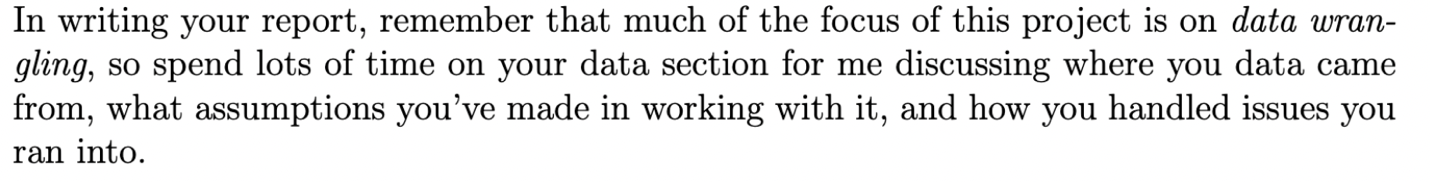
Chenxi Rong, Xiaoquan Liu, Fabian Schmid, Zhanyi Lin

November 29, 2022

* Motivation for project [Motivation]
* Motivation for the research design being used [Motivation]
* Details of the data used and how different datasets have been related to one another

[May need adjustment by Fabian or Zhanyi]

Nick said: but I don’t know what you guys ran into, so I need your help!



* Summary statistics for your data [Zhanyi&Chenxi]
* Your analysis and your interpretation of that analysis [Zhanyi&Chenxi/waiting for graph]
* Conclusion [will add after finishing the analysis part by Xiaoquan]

**Motivation**

Opioids are a class of drugs that include prescription opioids (natural and semi-synthetic opioids and methadone), heroin, and synthetic opioids other than methadone (primarily fentanyl) that derive from, or mimic, natural substances found in the opium poppy plant and work in the brain to produce a variety of effects, including pain relief. Opioid drugs include prescription pain medicine and illegal drugs. Some people may experience euphoria, a joyful sensation of well-being, from opioids, whether they are legally prescribed or not. Opioids don't always generate euphoria, but for those who do, there's a chance they'll be used again and again because of how they feel. Therefore, even with a doctor's supervision, using opioids can pose risks. A person's tolerance and dependency to prescription drugs can develop over time, necessitating greater and more frequent dosages, finally leading to addiction and the person's turning to illegal markets in order to maintain their addiction, and subsequently causing death.

According to the Centers for Disease Control and Prevention (CDC), the number of drug overdose deaths has quintupled since 1999, and the rise in opioid overdose deaths can be outlined in three distinct waves: the first wave in the 1990s with increased prescribing of opioids, the second wave in 2010 with heroin, and the third wave in 2013 with synthetic opioids like fentanyl. In order to fight the opioid overdose epidemic, policymakers have made policy interventions to limit the over prescription of opioids. Texas regulations with regard to treating pain with controlled substances went into effect in January 2007. Florida’s legislature became effective in 2010, and a series of changes related to drug prescription took place in the following years. Washington regulated the prescribing requirements of opioids for pain treatment in January 2012, which included periodic patient reviews, milligram thresholds, strict documentation guidelines, and consultations with pain management experts.

For all three of these policy changes, we performed both pre-post analysis and difference-in-difference analysis to understand the effect of opioid drug regulations on both the amount of opioid shipments and drug overdose deaths. For pre-post analysis, we will demonstrate the trend of overdose deaths and opioid shipments over years. If policy had gone into effect, our plots would show a difference between how things were in each state right before the policy went into effect and right after the policy went into effect.

However, to further valid our analysis of causation between opioid drug regulations and both the amount of opioids shipments and drug overdose deaths, we need to eliminate the effect of confounders. For example, the US Customs Service managed to dramatically reduce the importation of fentanyl into the United States at the same time Florida’s policy went into effect, which would likely reduce the number of overdose deaths throughout the United States. If we were just to use pre-post analysis to bring a conclusion by comparing Florida in 2009 to Florida in 2011, we would wrongly attribute the decline in the amount of shipments and overdose deaths to Florida’s policy change. With difference-in-difference analysis, we used the observed outcomes of people who were exposed to drug regulations (i.e., data from Texas, Florida, and Washington) and people who were not exposed to drug regulations (i.e., for each of those states, we picked three states as comparison states) both before and after the policy went into effect to evaluate the impact of opioid control policies.

**Data:**

We used drug overdose death data from the US Vital Statistics records, prescription opioid drug shipments from the Washington Post, FIPS codes based on a file from the US Census, and US census population data. All these data sets need at least a location name (to infer the county) and a temporal unit (to infer the year). The data sets were merged based on the county FIPS codes and the year. FIPS codes are already included in the population data set and the US Vital Statistics records. Besides that, the raw data was aggregated at the county-year level so that the data was available for our preferred unit of observation.

For the opioid shipment data, we used the total active weight of the drug in grams and converted it to a morphine equivalent so that different kind of opioids are comparable. To deal with the huge raw data set (> 100 GB), we only imported the required columns, chunked the data, and grouped it by year and county to reduce the number of rows and end up with our preferred unit of observation. Unfortunately, this data set only contains the county name but no FIPS code. The merging process based on the county and state name, required the harmonization of the county names in both data sets. For that, several steps were taken like capitalizing all letters and dropping suffixes. To check if the merging process was successful after the modifications, a test was run with the assert statement.

For the mortality data, we assigned the following five death causes as drug related: "Drug poisonings (overdose) Undetermined (Y10-Y14)", "Drug poisonings (overdose) Unintentional (X40-X44)", "Drug poisonings (overdose) Suicide (X60-X64)", "Drug poisonings (overdose) Homicide (X85)", and "All other drug-induced causes". Again the FIPS code caused problems in this data set. Although they were already included, some of them had only four digits so that it was required to add a zero to the beginning in these cases so that the merging later on works without problems.

The US census population data is published once every decade. As we require data for two decades, we used two datasets and concatenated them at the end. The population data set also required modifications to the FIPS code as it only included the two-digit state FIPS code and the three-digit county FIPS code. To combine them, it was again necessary to add zeros in the beginning if they have only one or two digits. Furthermore, it was necessary to reshape the data from wide to long so that we have our preferred unit of observation.

Based on our research questions and data collected, we merged the data into two master sheets. The first one contains opioid shipment and population to analyze how the policy influences the opioid convert amount. The second dataframe contains the overdose death data and population data.

As for the master sheet for opioid convert data and population data, it is merged with two steps. Firstly, because the opioid data contains only county name and state name, but population data only has FIPS code, we firstly merge opioid convert amount dataframe with FIPS dataframe. The merging process generated some opioid convert amount dataframe only observations. With indicators, we found the reasons related to the different spelling of County Name in FIPS data and in shipment data. After unifying the spelling of county name in FIPS and opioid shipment data, the merge succeed with pattern of m:1. Then we merge the intermediate dataframe with the population data based on *FIPS* and *year* for both dataframe since the data in shipment data and population data are both grouped by *FIPS* and *year.* After the merge, we found there are shipment data only observations. We checked the specific counties and states which miss the population data. The output indicated that the counties which have no population data are from PR (Puerto Rico) which is a Caribbean Island and unincorporated U.S. territory island. This state is not our target analysis state or in our selected control group. Hence, we safely drop those counties.

As for the master sheet for overdose death and population, since both dataframe has the features of *FIPS* and *year* code, we directly merge them based on those two features. We also employed *indicator* and *validate* to check the merge results. Firstly, the *validate* showed that the observations are merged in m:1 pattern. However, the indicator returns that in addition to AK, there are still two counties in VA are missing with population data. We examined the population data by FIPS of those two counties and found the population data doesn’t include those two counties. As we discussed in former part, Virginia is also needed in our analysis. Hence, we currently drop those observations. In our further discussion and analysis, if those counties are needed for future analysis, we will use Google to find the corresponding population of those two counties.