Project Proposal

Nathan Rosenfeld

November 19, 2021

1 Background

One of the most significant developments in the political landscape here in the United States over the last 20 years has been the shifting of the white working-class vote towards the Republican party. This shift seems especially prominent in the rural Midwest and Great Plains states. Another development in the realm of public health - a catastrophic one - has been the opioid epidemic.

Some political observers have remarked, at least on the basis of anecdotal evidence, that there may be some correlation between these two phenomena. Some have suggested that the crisis of drug addiction has caused people to blame welfare programs for, in their mind, enabling the addictions, thus driving people who were previously supportive of the social safety net to oppose it.

Others have suggested that some economic malaise caused by globalism is simultaneously driving both shifts. The purpose of this project is to determine whether there is any statistical correlation between these two occurrences.

2 Idea

The idea is to determine whether the extent of the opioid crisis correlates with the shift in political sympathies. In this project, I intend to use statistical regression to determine whether, to what extent, and within what confidence intervals the rate of increase of overdose deaths can be used as a predictor of the shift of the two-party vote.

3 Data Sources

- `https://data.cdc.gov/NCHS/NCHS-Drug-Poisoning-Mortality-by-County-United-Sta/pbkm-d27e`
  Contains number of drug overdoses by county from 1999 to 2015.
- `https://data.cdc.gov/NCHS/VSRR-Provisional-Drug-Overdose-Death-Counts/xkb8-kh2a`
  Contains the provisional count for overdose deaths by county in 2020.
- `https://electionlab.mit.edu/data`
  Contains the results of Presidential and Senate election broken down by county for the years 1976-2020.

4 Proposed Method

The proposed idea is to determine if there is a correlation between the growth in overdose deaths and two-party shift in partisanship. To do this, the following steps will be taken to analyze the data:

1. The data must be cleaned up.
   - In counties where the total numbers are very low, the number is masked in the dataset so as to protect privacy. Counties with results so masked should be left out of the analysis.

- The two datasets have to be given matching code numbers to identify the counties, to match the election data with the CDC data.

2. For each county, a regression line of overdoses as a function of time should be determined.
3. Then, another regression line of partisanship as a function of time should be determined for each county.
    - This "smooths out" the inherent "noise" of each election cycle (the respective strength of individual candidates, the news cycle at that moment, etc.)
4. Finally, a regression line of partisan-shift vs. overdose increases should be determined.
5. Confidence intervals and the power of the test should be determined and reported.

5 Limitations

There will be some inherent limitation in data that is broken up by county. If CDC data were reported by precinct, the increased granularity would make the study better. I'm not aware of any such data.