Read Me

Kristopher Velasco and Sebastian Rojas Cabal

2/25/2022

Last update by: Sebastian

For getting the 2008-2013 data

1. Get PDFs from Citizen Audit. Kristopher has the code to do that.
2. Figure out the code or program we are using for extracting the information.

- Layout parser: https://layout-parser.github.io/
- Company Kristopher is contacting.
- Layout parser: https://layout-parser.github.io/

3. Collecting the first page (Part I) of Form 990, and Schedule F (Part I and Part II; pp. 1-2)
4. We will use cases up to 2013, to avoid anticipation effects coming from the repeal of DOMA. USE THE CODE ON GITHUB TO GET THE STUFF FROM CITIZEN AUDIT. BY 3 PM WE WILL KNOW WHATS UP.

Tuesday, May 31 - Friday, June 3

- **THIS IS PAUSED UNTIL WE GET ALL THE DATA FROM 2008-2013, FROM OCR**
- New samples for descriptive figures.
- **FOR EASE OF DOWNLOAD AND DATA MANAGEMENT** the sample is divided into anti-LGBTQ+ nonprofits in one file and non-anti nonprofits in another one. They are named as follows:
  - **anti_sample_220601.csv**.
  - **nonanti_sample_220601.csv**.
- Both data sets include a variable for foreign and total expenses in constant 2013 USD.
- Both data sets are restricted for 2013-2020
- Only the **anti** data set includes a variable for the year in which marriage equality became legalized in the state.

Monday, May 16 - Friday, May 20

- Created a new ReadMe file to keep track of changes made to the project.
- Worked on the first draft of code for implementing SCM with staggered adoption. See below.
- **MAIN TAKEAWAY FROM MODEL:** Code for the model has been written. Implementation should be simple yet, we cannot run the model right now because we do not have enough expense data previous to the policy change. Our data begins in 2013 and most states passed legislation around 2012.

First draft of code for implementing synthetic controls with staggered adoption.

- Final sample for SCM consists of 608 organizations. These are organizations for which we have expense data continously from 2014 to 2018.
- Analytical sample consists of 1,144 organizations (unique EINs in the data). I removed the following observations. File available via One Drive. Name: **scm_sample_draft_220520.csv**.
    - All cases when Functional Expenses were 0.
    - Keeping only the most recent tax filing for each tax year, according to the return time stamp.
- Original sample consists of 2,061 organizations. These are known anti-LGBTQ+ nonprofits as well as organizations that have received money from them according to Schedule I data. Duplicated EINs and names were dropped. More info below.
    - 620 known anti-LGBTQ+ organizations. I removed duplicate names and EINs from the original list of 657 organizations.
    - 1,446 presumed anti-LGBTQ+ organizations. These are orgs that received money from known anti-LGBTQ nonprofits. I removed duplicate names and EINs from the original list of 1,795 organizations.
- The analytical sample merges four sources of data:
    - Form 990 return headers, parts 0 and part 9.
    - The list of candidate anti-LGBTQ+ orgs described above.
    - The **sample is restricted only to anti-LGBTQ+ orgs**. Synthetic control methods compare orgs to themselves before the policy change.
- Year of adoption was year when state legislature legalized marriage equality.
- Outcome is real 2013 dollars.
- Tax year: end date of tax year in the tax return.
- I chose the year marriage (not civil unions) were legalized. Data from: https://www.lgbtmap.org/ equality-maps/marriage_relationship_laws

Recap about where samples come from.

We used the information on the Schedule I forms we have to come up with:

- We had an initial list of 657 known anti-LGBTQ+ orgs Kristopher built based on his subject-area expertise. There are some duplicate EINs and names in this list.
- List of 1,795 first-order ties. These are orgs that have received money from known anti-LGBTQ+ orgs according to all our Schedule I forms. This file is called **anti_lgbtq_candidates_20220516.csv**. There are some duplicate EINs and names in this list.
- List of 162,547 second-order ties. These are orgs that have received money from orgs that received money from known anti-LGBTQ+ orgs according to all our Schedule I forms. This file is called **anti_lgbtq_candidates_20220422.csv**. There are some duplicate EINs and names in this list.## Monday, May 16
- Created a new ReadMe file to keep track of changes made to the project.