P395AnalysisDraft

Jasir Rahman

2023-11-07

Setup:

##11/04/23  
rm(list=ls())   
#setwd  
setwd("~/Desktop/R/POLI 395/POLI 395 Project")  
#libraries + packages  
library(devtools)

## Loading required package: usethis

library(foreign)  
library(qss)  
library(ggplot2)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(stargazer)

##   
## Please cite as:

## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.

## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

library(stringr)  
library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ forcats 1.0.0 ✔ readr 2.1.4  
## ✔ lubridate 1.9.2 ✔ tibble 3.2.1  
## ✔ purrr 1.0.2 ✔ tidyr 1.3.0

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

#Load in data  
og <- read.csv('original\_data.csv')  
blob <- read.csv('bills\_lobbied\_on.csv')   
lobby <- read.csv('lobbying\_data\_full\_2020.csv') #lobbying data  
leg\_116 <- read.csv('house\_legislation\_116.csv') #House legislative data  
mem\_116 <- read.csv('house\_members\_116.csv') #House member data  
bills <- read.csv('bills\_116.csv') #House and Senate legislative data

Data cleaning:

#set filtering strings  
filter\_string\_lob = c('Firearms/Guns/Ammunition','Firearm','firearm','Gun','gun')  
#join on id.poli395  
blog <- blob %>%   
 full\_join(og, by = c('id.poli395' = 'id.poli395'))  
#filter data to include only gun violence related legislation  
blog <- blog %>%   
 filter(str\_detect(specific\_issues, paste(filter\_string\_lob, collapse = "|")))  
blog <- blog %>%   
 filter(str\_detect(general\_issues, paste(filter\_string\_lob, collapse = "|")))   
#restrict to the 116th congress  
blog <- blog %>%   
 filter(congress == 116)  
  
#subset data to include only necessary columns  
blog = blog[, c(1:5,8,10,17,18,20,27,28,30,31,38)]  
#rename H.R. in data  
blog$bill\_name <- gsub("H.R. ", "H.R.", blog$bill\_name)  
  
#assign bills to anti-gun stance (for ease of computation)  
blog$billstance = rep('restrict', nrow(blog))  
#assign bills w/ pro-gun stance to list  
pro\_gun\_filter = c(110,121,155,1761,2179,2443,38,3826,5289,5935,6126,664,69,7715,775,817,877,821)  
#assign bills w/ neutral stance to list  
neutral\_filter = c(1222,1767,2075,2457,2698,3742,7614,7667,94,2492,286,5469)  
  
blog$billstance <- ifelse(blog$bill\_number %in% pro\_gun\_filter, 'expand', blog$billstance)  
blog$billstance <- ifelse(blog$bill\_number %in% neutral\_filter, 'neutral', blog$billstance)  
  
##check laws to see stances towards firearms  
blog$lobbystance = rep('restrict', nrow(blog))  
##check orgs manually to see general stances towards firearms  
#create filter list for pro-gun orgs  
pro\_org\_filter = c('Smith And Wesson Brands Inc.', 'Sig Sauer Inc', 'Shotspotter Inc.','Polymer80, Inc.','Palmetto State Armory, Llc',   
 'Olin Corporation', 'Nst Global, Llc', 'National Shooting Sports Foundation', 'National Rifle Association Of America',   
 'National Rifle Association, Institute For Legislative Action', 'Nst Global, Llc (Dba Sb Tactical)',  
 'National Association For Gun Rights', 'Nammo Perry, Inc', 'Mars, Inc.', 'Kongsberg Defence & Aerospace', 'Kelley Drye & Warren Llp',  
 'Hill Country Class 3, Llc', 'Gun Owners Of America Inc', 'Firearms Regulatory Accountability Coalition, Inc. - Frac',   
 'Firearms Regulatory Accountability Coalition, Inc. (Frac)', 'Firearms Policy Coalition, Inc.', 'Defense Distributed',   
 'Dallas Safari Club', 'Corporation For The Promotion Of Rifle Practice & Firearm Safety',   
 'Citizens Committee For The Right To Keep And Bear Arms', 'American Outdoor Brands Corporation', 'Magpul Industries Corp.')  
#create filter list for neutral/mixed orgs  
neutral\_org\_filter = c("National Fraternal Order Of Police", "Dick'S Sporting Goods (On Behalf Of Finsbury Llc)")  
  
#assign rows of pro-gun orgs label 'expand'  
blog$lobbystance <- ifelse(blog$client\_name %in% pro\_org\_filter, 'expand', blog$lobbystance)  
#assign rows of neutral orgs label 'neutral'  
blog$lobbystance <- ifelse(blog$client\_name %in% neutral\_org\_filter, 'neutral', blog$lobbystance)  
  
##merging bill statuses on bill names'  
bills = bills[, c(3,4,5,7,8,9,10,12)]  
bills$bill\_number <- gsub("HB", "H.R.", bills$bill\_number)  
bills <- bills %>%   
 rename(bill\_name = bill\_number)  
  
#rename variable name to sponsor in mems\_116  
mem\_116 <- mem\_116 %>%  
 rename(sponsor = name\_id)  
  
#create list of strings from all bill names   
leg\_116 = leg\_116[, c(1,3,4)]  
  
  
#merge data based on column sponsor for both bills  
sponsor <- leg\_116 %>%  
 full\_join(mem\_116, by = c("sponsor" = "sponsor"))  
  
#filter data by keeping inclusive to cells in bill\_id with data  
sponsor <- sponsor[complete.cases(sponsor$bill\_id),]  
sponsor <- sponsor %>%   
 rename(bill\_name= bill\_id)

Merging datasets:

#merge data  
data <- inner\_join(blog, sponsor, by = "bill\_name")  
data <- inner\_join(data, bills, by = "bill\_name")  
#filter out data for bills without gvp focus  
data <- data %>%   
 filter(str\_detect(description, paste(filter\_string\_lob, collapse = "|")))

Checking summary statistics:

#check on summary statistics for ind. or dep. variables  
summary(data$dollar\_amount)

## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
## 10000 80000 100000 211252 320000 1400000 76

summary(data$status)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 1.000 1.000 1.107 1.000 2.000

#check and print percent missing for ind. and dep. variables  
missing\_percent\_dollar\_amount <- sum(is.na(data$dollar\_amount)) / length(data$dollar\_amount) \* 100  
missing\_percent\_status <- sum(is.na(data$status)) / length(data$status) \* 100  
cat("Percentage of missing values in 'dollar\_amount':", missing\_percent\_dollar\_amount, "%\n")

## Percentage of missing values in 'dollar\_amount': 5.946792 %

cat("Percentage of missing values in 'status':", missing\_percent\_status, "%\n")

## Percentage of missing values in 'status': 0 %