ECON 6750 Data Replication Project
By Jacob Croskey

Paper Details:
- Title: "Sticking to One's Guns: Mass Shootings and the Political Economy of Gun Control in the United States"
- Author: Dr. Hasin Yousaf
- Published: March 30, 2021
- Link: https://doi.org/10.1093/jeea/jvab013

Write-Up:

The causal relationship of interest in this paper is on the impact of mass shootings on voter behavior in the counties where they occur. More specifically, the paper analyzes the effects on Republican vote share and aims to investigate the political preferences associated with policy issues such as gun control and gun violence. The ideal experiment for this relationship would be a randomized controlled trial where mass shootings are randomly assigned to different counties and then the Republican vote share is analyzed. However, the practical and ethical challenges associated with an experiment like this would be too difficult to overcome, so the author uses a difference-in-differences strategy for identification. This strategy compares the vote share between similar counties where some counties were affected by mass shootings while others were not over the same time period. This allows us to isolate this causal effect.

The title of the paper in the literature review that seemed closest to this paper is called, *"Reexamining the Effect of Mass Shootings on Public Support for Gun Control"* [1]. This paper focuses on the broader implications of mass shootings and connects it tightly with the effect on gun control regulations. It uses panel data, controlling for geography, in order to see how opinions on gun control laws change over two years based on whether or not a person lived near a shooting in that time span. I think this paper improves on this idea because it takes the public opinion aspect of the earlier study and transforms it into the more tangible and digestible medium of vote share where we can really see the effect of these events. Furthermore, the vote share analysis bridges the gap between the public opinion of how these mass shootings make people feel and how those feelings are translated into political choices when a person enters the voting booth. The second, and more recent, paper is called, *The Effects of Mass Shootings on the Political Behavior of Victims' Families and Neighbors* [2]. The newer paper improves on this paper by taking the analysis a step further and analyzing the voter and political change not only in the citizens of areas surrounding mass shootings but the family and friends that knew victims of these shootings. Simply put, it

takes the analysis to a more personal level. This paper analyzes the link between the emotional impact of mass shootings and the expression of those sentiments in the democratic process and political climate. I believe both the more recent article and the article I chose for this project have their pros and cons in terms of method of experimentation, but I do think the paper I chose does a better job at comparing causal effects with the use of difference-in-difference techniques as opposed to the panel data experiments in the other article which seems to have more individual-level insights.

[1] https://doi.org/10.1017/S0007123418000352

[2] https://www.proquest.com/dissertations-theses/effects-mass-shootings-on-political-behavior/docview/2700377537/se-2?accountid=14537

There are a few strengths and weaknesses in this paper. To start off the strengths, one of the most impressive things about the paper is the data it uses for its empirical analysis. It gathers data from the FBI along with multiple media sources to form a comprehensive list of mass shootings dating from 2000-2016 at a county level. Another strength to this paper is the use of both the difference-in-differences strategy as well as a comparison between failed and successful mass shootings. This gives the article more credibility by using two empirical analysis techniques while also diversifying investigation techniques to provide a better picture of the causal effect. The final strength I found in this paper relates to the consideration of multiple election outcomes. It includes House, Senatorial, and Gubernatorial elections which contributes to a better understanding of how and by how much the political implications of mass shootings are changing. However, there are some weaknesses that should be considered when reading this paper. The first I noticed was the assumption that the success or failure of a mass shooting was exogenous, meaning it is not influenced by factors not controlled for in the analysis. The paper does attempt to justify this assumption, but the randomness of failure or success of shootings may introduce some bias into the estimated effects. Examples of unobserved factors may include law enforcement response timing or preventative security measures. The second potential weakness in the paper stems from the focus of the paper on polarization rather than taking a closer look at the moderate opinion. The paper does a good job of analyzing the polarized opinion shifts, but could do a more thorough analysis of the changes in opinion in the middle ground of the debates. Understanding how mass shootings affect more moderate citizens could have more, and possibly better, information on the true causal effect and overall impact.

I had a little bit of trouble creating a nice table in R for the estimates, so I have provided screenshots here for each replication I performed. I have also included the entire code later in this document for a more complete look at the processes I used (I cut several pages that just had hundreds of "stateyear" interactions to reduce page

count). Additionally, I provided short summaries of each regression I ran and how they were different from one another in the full code below.

Table 4 Panel A Estimates

```
      mspost
-0.06166999
      mspost
-0.03733891
      mspost
-0.02984154
      mspost
-0.04307539
      mspost
-0.03940394
```

Reading from top (Model1) to bottom (Model5) we can see the effects of mass shootings on Republican vote share. Model1 is the simple relationship without controlling for state or population time trends. Model2 includes the state and population time trends. Model3 includes state-year fixed effects. Model4 and Model5 are the same as 2 and 3 respectively, but are weighted by county population. Most of my estimates were spot on with the exception of one or two which I attribute to the difficulty in setting fixed effects (described more below).

Table 4 Panel B Estimates

```
 successful   postattack
-0.03459583   0.00918278
 successful    postattack
-0.031960055  0.007528018
 successful   postattack
-0.03197771   0.00753166
 successful   postattack
-0.02311337   0.02043918
 successful    postattack
-0.019178521  0.004355988
```

Once again, reading from top (Reg1) to bottom (Reg5) we can see the effects of successful mass shootings in Republican vote share. The "postattack" acts almost like a control to show that unsuccessful attacks have no (or negligible) effect on Republican vote share. Reg1 shows that successful mass shootings have an effect and failed do not. Reg2 introduces flexible population trends without changing coefficients much.

Reg3 introduces state-year fixed effects. Reg4 and Reg5 are similar to those in 2 and 3 respectively. Most of my estimates were spot on with the exception of one or two which I attribute to the difficulty in setting fixed effects (described more below).

To elaborate on the methods used in this study, I conducted a collinearity check among the independent variables. Generating the correlation matrix above allowed me to examine the relationships between these variables, aiding in the identification of potential issues arising from high correlation. Apart from the "state" and "stateyear" correlation which is expected because they derive from the same thing, no other independent variables are close to 1 which supports the idea of perfect collinearity.

	state	year	county	pop	stateyear
state	1.000000000	0.0056701417	-0.0038811221	-0.05144197	0.998686355
year	0.005670142	1.0000000000	0.0004485264	0.01807082	0.025354195
county	-0.003881122	0.0004485264	1.0000000000	0.01231381	-0.003128657
pop	-0.051441973	0.0180708249	0.0123138106	1.00000000	-0.047128701
stateyear	0.998686355	0.0253541954	-0.0031286570	-0.04712870	1.000000000

The replication for this project was quite hard to do. While the code was well-organized, there were little to no comments in the code describing what the code was doing which made it difficult to understand what each command and function was doing with the data. Additionally, while the author did mention which characteristics he changed for each new column in his panels, he failed to describe which variables he used in the regressions and what variables were used as part of the fixed effects. For example, when running the regressions and including population and state fixed effects, the author did not mention whether he used "pop", "popsd", "popsdsq", or "lnpop" which led to a guess and check method for trying to replicate the estimates. The author published more complete datasets than was needed to replicate the results, but some of the variables did not seem necessary to include and seemed to simply take up space.