School Shootings

A Statistical Investigation into Firearm Safety Scorecards

Introduction

Gun violence is an ongoing and concerning threat to the American people. Recent solutions to reducing gun violence target legislative actions for increased controls on the transfer, sale, and possession of firearms. The implementation of these actions is mired in conflict ranging from personal moral beliefs to Constitutionally-inalienable rights.

Mass school shootings are an emotionally charged subject, which have promoted legislative gun control measures with wideband authoritative powers encompassing mental health, right-to-buy, and possession requirements. It's generally accepted that any solution to gun violence requires consideration of an exhaustive set of environmental, political, and ethical variables; meanwhile, the challenge and debate remains divided on which variables deserve actionable consideration.

The intent of this report is to determine if there exists a relationship between school shootings, comprehensive [State] firearm safety laws, and firearm deaths across the United States.

Literature Review

This study focused on CNN Lab's curated school shooting data as its primary source of event observations. Data from the Centers for Disease Control (CDC), National Death Index (NDI) and the US Census Bureau provided supplemental population data to establish comparative values surrounding firearm deaths.

Comprehensive firearm legislation scores were sourced from Giffords Law, an anti-gun organization who has prepared State-level scorecards for over a decade. Their scorecards have consistent methodologies and are considered a reputable standard in the political space of gun lobby and legislation. Information from their annual scorecard reports was be used in the study as the primary comparison to CNN Lab's event data.

Methodology

School shootings were identified by the severity [count] of injuries, school location, and various demographics about the shooting's epicenter. Giffords' state firearm legislation scores are ordinal, as prepared by Giffords' methodology. Source information from NDI and US Census Bureau was used to independently validate information

reported in Giffords' Scorecards. This study focused on categorical analysis and primarily considered the Chi-squared test for independence until distributions challenge the efficacy of the test. This study evaluated the low efficacy of the Chi-squared test was due to low observational data, and proceeded to use Monte Carlo simulations and Fisher Exact tests as alternatives.[i]

Findings

This study discovered four key results from its analysis:

1. Giffords Gun Scorecards show no dependence or relationship between State gun law ranks and school shooting events.
2. Giffords Gun Scorecards show no dependence or relationship between State gun law ranks and school shooting victims.
3. School shootings occurred more often in States with moderate firearm death rates, as opposed to low or high rates.
4. States with the highest firearm death rates do not represent the highest risk for school shootings.

The combined results of this study provided us the conclusion that Giffords Gun Scorecards do not represent a valid composition of firearm safety ratings with regard to the subset of school shootings, nor do current firearm rights or controls have much effect on school shooting dangers. This study assumed Giffords' evaluation was a suitable measure of gun violence and firearm risk; our independent analysis confirmed key elements of this assumption, in line with Giffords' publications. However, having results be contrary to this assumption is unintuitive, and reflects on the complexity of how firearm laws, safety, and control measures are implemented in our society.

This study looked to existing research to identify school shooting motivations in order to determine possible reasons for its findings. The US Government Accountability Office (GAO) identified statistics with curious overlays in their 2020 Congressional report, stating:

1. School shootings were conducted by current or former students half the time of all incidents.
2. Incidents were the school (or persons at the school) approximately 30% of the time, and school grounds were incidental or unintended locations almost 36% of the time.

Analysis

Data Overlay

This study utilized CNN Lab's collective event-based data on 180 school shootings from 2009 – 2018. The distribution of data is shown below for reference. The dataset identified the severity of the shooting by counting the number of injured survivors and number of decedents (collectively referred to as *victims*) for each event. Additionally, location and demographic information was provided, however this study did not focus on these fields due to low utility with the demographics provided, and scope of the study.

Event Timeline Range	2009 Jan 08 – 2018 Dec 11	
Total Number of Events	180	
Distribution of Victims/Event	Range	1 – 31
	Mean	1.9778
	Median	1
	Total	356
Distribution of Deaths/Event	Range	0 – 26
	Mean	0.6333
	Median	0
	Total	114 (32.0 % of Victims)
Event Locations (# Deaths/ # Victims)	Elementary School	30 (39 / 67)
	Middle School	24 (6 / 31)
	High School	118 (67 / 245)
	Combined Levels	8 (2 / 13)

NDI multiple cause mortality data compiled per state per year was used to independently support information from Giffords' publications and provide strata overlays into the analysis of this report.[ii] NDI information was joined into US Census Bureau results to create population baselines so firearm death rates can be calculated. Death rates were calculated in the normal unweighted fashion commonly presented by the CDC, and shown below for calculation reference.

$$d = \frac{\text{total deaths}}{\text{total population}} \times 100{,}000$$

NDI data was classified into groups to aide analysis, and multiple bands of classification and binning were used to determine if results changed. Binning was restricted to near-linear equity to prevent bias from non-linear distributions. This report identified no viable change in its results given different linear bins.

Distribution of Firearm Deaths	Range	23 – 3,622
	Mean	692
	Median	505
Distribution of Population	Range	563,626 – 39,309,796
	Mean	6,366,419
	Median	4511281
Distribution of State Death Rates	Range	1.6 – 23.5
	Mean	11.7
	Median	11.7
Classification of Death Rates (bin)	Low	1 – 6.9
	Moderate	7 – 13.9
	High	14 – 23.9

Giffords Scorecard Ranks

This study tested the relationship between Giffords' Scorecard Ranks and CNN Lab's events and victims. Utilizing Chi-squared tests, this report showed no evidence of independence between state's Gun Law Rank and the number of victims, resulting in a test p-value of 0.3577. Conversely, the same test structure in comparing state's Gun Law Rank and the number of shooting events did suggest independence, resulting in a test p-value of 0.0358.

Given the populations of each element in the confusion matrix of these tests, this study believed it prudent to validate the integrity of these Chi-squared findings. Utilizing Monto Carlo simulation, the source populations were randomly selected 10,000 times to determine the populations elemental to the Chi-squared test statistic. Furthermore, each test was repeated with the Monte Carlo simulations 1,000 times to identify the distribution of the test statistic and resulting p-value. This study intended to question the legitimacy of any test result where the simulated distribution approached the 0.05 alpha-level threshold.

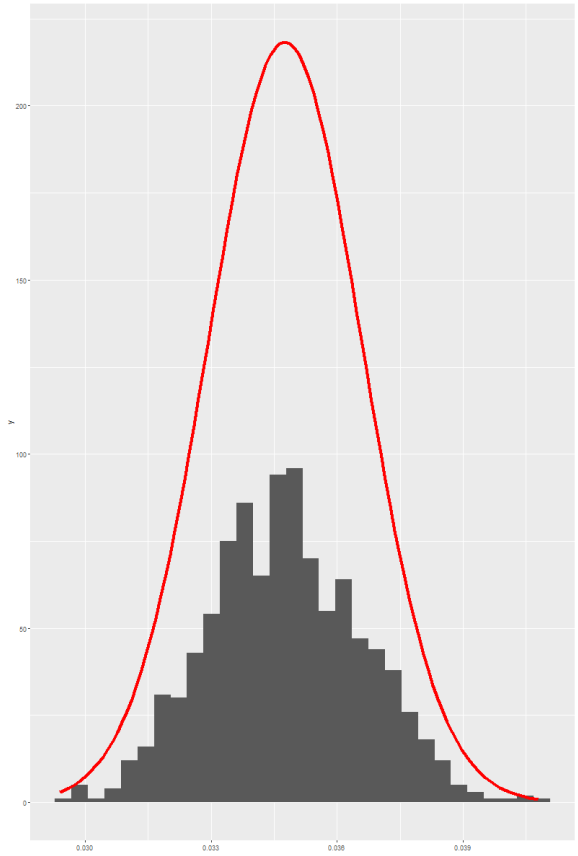

FIGURE 1: P-VALUE SIMULATION OF GUN SCORECARD RANKS VS. EVENTS

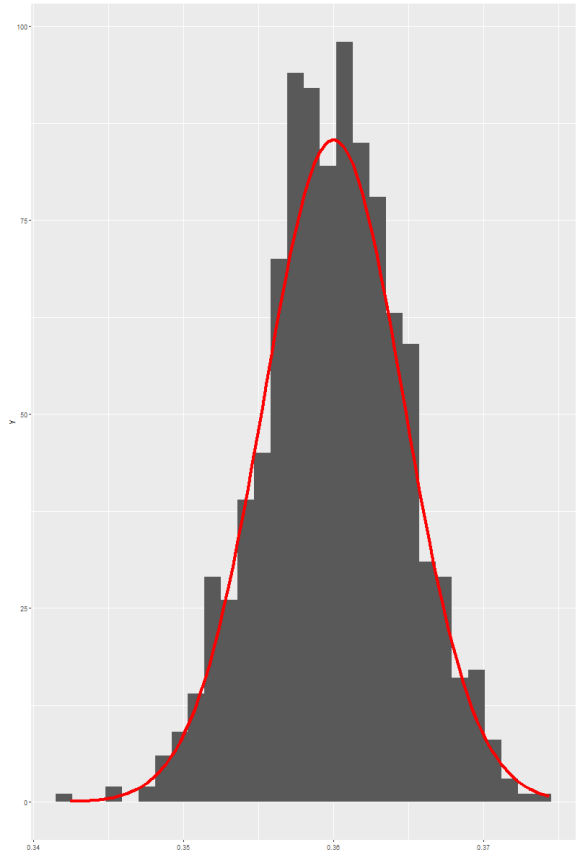

FIGURE 2: P-VALUE SIMULATION OF GUN SCORECARD RANKS VS. VICTIMS

Here we notice the distributions are not both normally distributed, signifying a potential underfit to the model evaluating Gun Scorecard Ranks versus Events. This suggested Chi-squared was an incompatible statistic to use for this analysis. Due to the underfitting of the simulation, along with the weakness of Chi-squared tests when expected populations are very low, this study resorted to Fisher Exact tests for better evaluations. With the same variables as above, this study found Fisher Exact tests to report no evidence of dependence between Giffords' Scorecard Ranks and total Events, or Victims; the final p-values for the tests were 0.3299 and 0.3789, respectively.

Given the results, this study rejects its initial findings with the Chi-squared test, and finds there to be no discernable relationship with Giffords' Scorecard Ranks and either school shooting Victims or Events.

Firearm Death Rates

Giffords' Scorecard Ranks were plotted against calculated Death Rates for each state, and results compared against similar [Giffords] graphs published. This study found the trends comparable, with a concern Giffords' death rates were consistently higher than direct calculations done in this study. The differences ranged from approximately 1-3 deaths per 100,000. This study suggests further investigation on the differences.[iii]

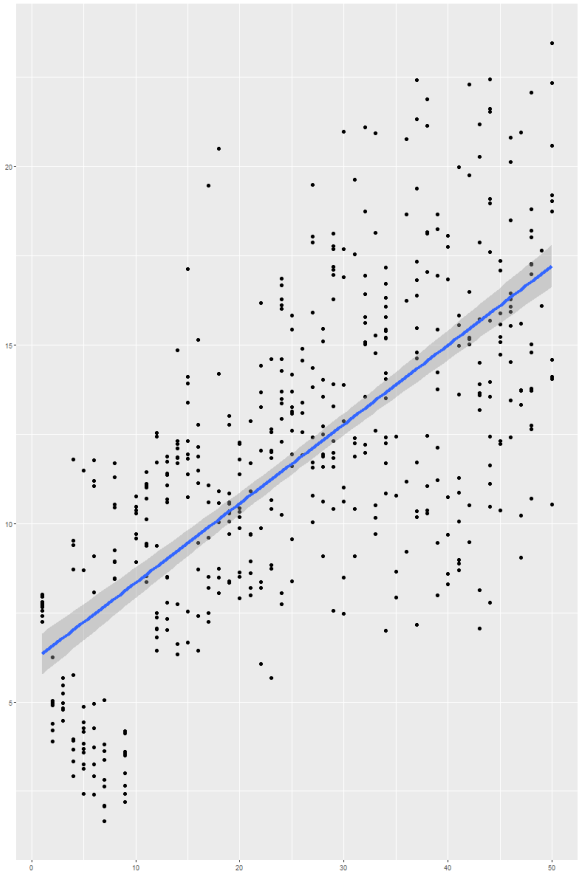

FIGURE 3: GUN SCORECARD RANKS VS. FIREARM DEATH RATES

Applying Fisher Exact tests to Death Rates and States' Victims yielded highly significant findings. With a p-value of 0.0029, this study finds there is a significant dependence between States with moderate firearm deaths, and total victims. This study expected the propensity of victims and violence to increase as total firearm rates increased, however, these results inform us it is a nonlinear trend, having some plateau or peak effect of victims as total firearm deaths increase.

Given these results, this study rejects the notion an increase in school shooting events and victims directly correlate with an increase in firearm deaths. Instead, this study believes there is an undetermined interval where the correlation is strong and positive, and a second [unknown] interval where the correlation is less strong and negative.

Count Row %	Total State Firearm Death Rate Low	 Moderate	 High	
No Victims	64 16.4	212 54.4	114 36.2	**390**
Yes Victims	5 4.5	68 61.8	37 33.6	**110**
	69	**280**	**151**	**500**

Appendix

Data Caveats

Data utilized in this report were evaluated for limitations, completeness, and variability which may have negatively impacted the methodology and findings of this report. All considerations are annotated in detail here, for reference.

Giffords Gun Law Scorecards

Annual reports for Giffords Gun Scorecards were scraped from their source website into a consumable format. Annual report information for years 2010, 2013 – 2019 were obtained directly, without transformation. Scorecard information was unavailable for years 2011 and 2012 due to missing publication information from Giffords. This study chose to impute the values of 2011 and 2012 [gun law] Scorecard Ranks to prevent missing information from impeding reasonable analysis and results. In order to maintain consistency of information from known years 2010 and 2013, this study selected on a linear calculation for each state rank, such that the net change is as equally-balanced as possible, with no skew from 2011 – 2013, and a net change as close to the intra-year integer minimum as possible. The calculation of the imputed ranks is shown below, where s is each State's Gun Law Rank per subscripted year.

$$s_{2011} = s_{2010} + \text{floor}\left[\frac{(s_{2013} - s_{2010})}{3}\right]$$

$$s_{2012} = s_{2013} - \text{floor}\left[\frac{(s_{2013} - s_{2010})}{3}\right]$$

As a result of this imputed data, Gun Law Ranks for 2011 and 2012 are not ordinal unique and now distribute as a mixture of competition and modified competition ranks (Wikipedia 2023).

US Census Bureau

State population data was obtained from the official 2010 and 2020 Decennial Census reports. A linear projection between the two reports was used to determine the intra-decennial years' (2011 – 2019) populations, such that the change in population between consecutive years was constant. The formula is shown below for reference, where p_Y represents the computed state population for each subscripted year, Y.

$$p_Y = p_{2010} + \frac{(Y - 2010)(p_{2020} - p_{2010})}{10}$$

References

Centers for Disease Control and Prevention [CDC], National Center for Health Statistics [NCHS]. 2021. *National Vital Statistics System, Mortality, 1999-2020.* Accessed April 07, 2023. http://wonder.cdc.gov/mcd-icd10.html.

CNN Labs. 2019. *School Shooting Data.* July 26. Accessed April 06, 2023. https://github.com/cnnlabs/cnn-school-shooting-data.

Giffords. 2023. *Gun Scorecards.* April 06. Accessed April 06, 2023. https://giffords.org/lawcenter/resources/scorecard/.

Government Accountability Office. 2020. "K-12 Education: Characteristics of School Shootings." Accessed May 22, 2023. doi:GAO-20-455.

US Census Bureau. 2010. *US Decennial Census 2010, Table P1.* Accessed May 10, 2023. https://data.census.gov/.

US Census Bureau. 2020. *US Decennial Census 2020, Table P1.* Accessed May 10, 2023. https://data.census.gov/.

Wikipedia. 2023. *Ranking.* February 08. Accessed April 22, 2023. https://en.wikipedia.org/wiki/Ranking.

[i] Chi-squared tests were evaluated with Monte Carlo simulation on the sample values to better approximate the population's critical values of the subject test. The resulting critical probabilities were replicated by secondary simulation to evaluate any bias in the Monte Carlo simulation. This study found the critical probabilities from Monte Carlo simulation produced non-normal distributions of data, in some cases.

[ii] Query Parameters: MCD - ICD-10 113 Cause List: Accidental discharge of firearms (W32-W34); Intentional self-harm (suicide) by discharge of firearms (X72-X74); Assault (homicide) by discharge of firearms (*U01.4,X93-X95); Discharge of firearms, undetermined intent (Y22-Y24).

[iii] This finding has no material effect on the analysis of this report as death rates are not formally evaluated or tested outside of linear binning for tests of independence.