

Temporal trends in public mass shootings: high-capacity magazines significantly significantly increase fatality counts, and are becoming more prevalent

December 12, 2019

Appendix A

Data

No complete and comprehensive official database of public mass shootings currently exists. Previous researchers have often relied on incomplete government databases, or databases compiled by unofficial sources such as USA Today and Mother Jones [1]. Unfortunately, like other researchers, we have found various compilations of public mass shootings to often be incomplete or have errors in the event details, and/or have apparent inconsistencies in criteria for inclusion [2, 3, 4, 5]. In this analysis we attempted to rectify this by compiling a list of all high-fatality public mass shootings in the United States from 1995 onward that satisfy a well-defined set of criteria. Here we define “mass murder” according to the 2010 Federal Bureau of Investigation (FBI) criteria of four or more killed [3], since it is these high-fatality count events which are most likely to receive widespread media attention, and foment debate on potential changes to state and federal policies to address the problem of firearm violence in the US [6, 7].

Events included in our analysis may have had multiple perpetrators, and may have occurred in multiple public locations, but must have occurred within a 24-hour period to be included. Shootings related to gang violence or other criminal activity were excluded, as were shootings that could plausibly have been incited for reasons of self-defense, and shootings that began with a police encounter during a traffic stop precipitated by suspicion of illegal activity. Only events with at least four people shot and killed in *public places*, not including the perpetrator, were included; victims killed during the incident on private property were not counted. We also only counted victims who died within one month of the incident, because it is the initial fatality count that primarily captures the public attention.

Sources of information used in this analysis included various print and online compilations of mass shootings (all URL’s accessed November, 2019):

- The FBI Uniform Crime Reporting Supplementary Homicide Reports
- Klarevas (2016) [2]

- Fox (2018) [8]
- Fox and Levin (2003) [9]
- Duwe (2007) [10]
- Krouse and Richardson (2015) [11]
- The USA Today online database of mass killings from 2006 to 2017, available at <https://bit.ly/1g9hHtz>.
- The Mother Jones public mass shootings database, 1982 to present, available at <https://bit.ly/2vYbqxF>.
- A tally of active shooter events in the US from 1996 to 2010 compiled by the New York City police department, available at <https://on.nyc.gov/2HHHI1N>.
- A tally of mass shooting incidents in the US involving large capacity magazines from 1984 to 2012, compiled by the Citizens Crime Commission of New York City, available at <https://bit.ly/2JGuL8B>
- The list of active shooter incidents from 2000 to 2012 compiled by the New York City Police Department (2012) [12]
- A compilation of public rampage killings with four or more people killed from 1940 to 2017, compiled by CNN, available at <https://cnn.it/2I2s7wK>.
- A list of public mass shootings with four or more people killed from 1966 to present, compiled with the Washington Post, available at <https://wapo.st/2LyZ1TF>.
- The list of suicide terrorists and rampage, school and workplace shooters from 1990 to 2010 compiled by Lankford (2013) [13].
- The list of mass shootings with three or more fatalities involving firearms with high-capacity magazines from 1981 to the end of 2017, compiled by the Violence Policy Center, available at <https://bit.ly/2rb6kJy>.
- The list of public mass shootings with three or more people shot (not necessarily killed), not including the shooter, and not identifiably gang, drug, or crime related, from 1966 to 2016 compiled by the Stanford Mass Shootings of America project <https://stanford.io/1HSVWel>.
- The list of mass shootings from January 2009 to July 2015 with four or more people shot and killed (not including the shooter), compiled by the Everytown for Gun Safety organization, available at <https://every.tw/2w0jI7V>.
- The list of active shooter events in the US between 2000 to 2016, compiled by the FBI, available at <https://bit.ly/2rb1s9R>.

- The list of mass shootings with four or more people shot and/or killed (not including the shooter), compiled by the Gun Violence Archive, which attempts to catalog every incident of gun violence in the US from 2013 to present, available at www.gunviolencearchive.org/.

In addition to these sources of data, we also sought data from the LexisNexis database, and online media reports and court documents, uncovered by the following Internet search terms: mass shooting, mass murder, mass killing, spree shooting, spree killing, rampage shooting, rampage killing, public shooting, and massacre. If, on occasion, media reports differed in their details regarding the number injured and killed, we took the most recent media report as the most accurate, and information available in court documents took precedence over information in media reports.

From these sources of data, events were selected that fit our definition of high-fatality public mass shootings.

There is significant overlap in many of the above listed compilations, but events were often included in some that were not included in others, in part due to differing inclusion criteria in the compilations [2]. However, the authors have noted several instances where events should have likely been included in some compilations (according to the stated event inclusion criteria) but were excluded without explanation as to why, and it is unclear whether the exclusion was deliberate, or instead due to ignorance of the incident on the part of the researchers compiling the list.

Accordingly, to avoid this confusion with our database, we created a secondary database of events that have been listed as “public mass shootings” in some past compilations, but did not meet the criteria for inclusion in our primary database (because, for example, all or most of the people killed were killed in private homes, not public places). An example of an event typically tallied as a “public mass shooting” by several past compilations, but was excluded in ours, is the September 2, 2008 incident in Alger, WA, where the perpetrator killed five people in private homes, then shot and killed one other during a subsequent public shooting spree. We included such events in our secondary database, with explanation as to the specific criteria each event did not meet for inclusion in the primary data. In our online visual analytics application, available at https://mass-shooting-analytics.shinyapps.io/mass_shootings/, we allow the user to repeat our analyses with the option to specifically include or exclude certain events. The visual analytics application also contains an extensive set of links to online material that were used as sources of information related to each event.

We selected incidents from the beginning of the FAWB on September 13, 1994 to the end of December 2018. However, because no incidents meeting our criteria occurred from September 13, 1994 to the end of 1994, the data selection effectively is for events between 1995 to 2018. We found 85 public mass shootings, 28 (32.9%) of which were stranger-on-stranger incidents.

For each incident, we examined media reports and other documentation to determine the number and types of weapons used, including whether weaponry banned during the FAWB was used, such as high-capacity magazines carrying more than ten rounds of ammunition (used in either a rifle or pistol), or an “assault rifle”. We use the definition of assault rifle specified in the FAWB legislation, which are semi-automatic rifles with a detachable

magazine that possess two or more of the following features; folding or telescoping stock, pistol grip, bayonet mount, flash suppressor or threaded barrel, or grenade launcher (see <https://bit.ly/2U6dv0S>, accessed November, 2019). For 90.6% incidents we were able to determine whether or not such weaponry was used.

Statistical methods

Temporal trends in frequency of incidents

Mass shooting events, per capita, are rare, and thus Ordinary or Generalized Least Squares methods are inappropriate for analyses of the trends in the frequency of such events [4]. Also, when analyzing the temporal trends in the frequency over a long time period, the effect of population change must be taken into account (for example, everything else being equal, doubling the population size will double the frequency). In addition, binning of data into a relatively long time period, like a year, necessitates loss of information and reduces sensitivity of the analysis [14]. In this analysis, we thus performed an unbinned likelihood fit to the data (beyond the necessary binning of data into integer days), and examined the temporal change in the per capita number of mass shootings per day, λ/N , where N is the population size, and λ is the average number per day. We employed a population-standardized log-linear model[15], with the predicted number per day, λ , described as

$$\log \lambda = \log N + \beta_0 + \beta_1 t, \quad (1)$$

where t is the date.

To account for potential over-dispersion in count data involved our analyses, we utilized a Negative Binomial likelihood fit [16]. The probability mass function (PMF) of the Negative Binomial distribution for observing k counts when λ are expected is [16]

$$f(k|\lambda, \alpha) = \frac{\Gamma(\alpha + k)}{k! \Gamma(\alpha)} \left(\frac{\lambda}{\lambda + \alpha} \right)^k \left(1 + \frac{\lambda}{\alpha} \right)^{-\alpha} \quad \lambda > 0, \alpha > 0, \quad (2)$$

where α is the over-dispersion parameter. The mean of the PMF is λ . When $\alpha \rightarrow \infty$ the Poisson distribution is obtained, and when $\alpha \rightarrow 0$ (i.e. highly over-dispersed data) the log-series distribution is obtained [16, 17, 18].

Given a set of M observations of the number of mass shootings per day, k_i , with $i = 1, \dots, M$ the likelihood of the observations is

$$\mathcal{L} = \prod_{i=1}^M f(k_i|\lambda(t_i), \alpha), \quad (3)$$

where $\lambda(t_i)$ is the expected number of counts on day t_i as determined from Equation 1. The best-fit model values used in the calculation of the $\lambda(t_i)$ are the values that maximize this likelihood [19].

Temporal trends in fatalities per incident

The number of fatalities per incident is integer count data, and Least Squares methods are again not appropriate for statistical analyses of such data because they are highly non-Normal

and skewed [4]. While some previous analyses have used Poisson regression methods to analyze mass shooting fatalities [20, 21], this does not take into account the over-dispersion in the distribution of fatalities, and can thus lead to erroneous over-estimation of the statistical power of the analysis. Negative Binomial likelihood methods are most appropriate for over-dispersed count data [16], but in an analysis like ours the model must account for the fact that the data are truncated by the selection of the minimum number killed in order to enter the sample (in the case of this analysis, four).

The PMF for Negative Binomially distributed data that are left-truncated at $k \geq 4$, $g(k|\rho, \alpha)$, is derived from the Negative Binomial probability mass function in Equation 2 as [22, 23]

$$g(k|\rho, \alpha) = \frac{f(k|\rho, \alpha)}{1 - f(0|\rho, \alpha) - f(1|\rho, \alpha) - f(2|\rho, \alpha) - f(3|\rho, \alpha)},$$

for $k \geq 4$, and is zero otherwise. In a similar fashion, left-truncated PMFs of the Poisson and log-series distributions can be obtained by substituting the appropriate PMF into Equation 4.

The log-linear model for the predicted number of fatalities for an event occurring at time t is

$$\log \rho = \beta_0 + \beta_1 t. \tag{4}$$

Given a set of M observations of the number of killed in mass shooting incidents, k_i , with $i = 1, \dots, M$ the likelihood of the observations is

$$\mathcal{L} = \prod_{i=1}^M g(k_i|\rho(t_i), \alpha). \tag{5}$$

Again, the best-fit values of the parameters used to calculate ρ maximize this likelihood.

It should be noted that, just like in the analysis of the temporal trends in the incidence over time, maximum statistical power of the analysis is attained when an unbinned likelihood fit is performed to incident-level data, rather than data aggregated over a time period like a year [14].

Comparison of the number of fatalities per incident for events that do, and do not, involve weaponry banned during the FAWB

Because mass shootings are rare, and the number of fatalities per incident are integer data truncated at $k \geq 4$ fatalities, hypothesis tests based on assumptions of Normality, such as the Student's t-test are not appropriate for comparing the means of the distributions [24]. The sample sizes are also generally not large enough such that the Central Limit Theorem applies, such that the means can be compared using the Z test. In addition, non-parametric tests that compare distributions, such as the Kolmogorov-Smirnov test, are not appropriate in the presence of ties (i.e. overlapping values in the data sets) [25].

Here we thus relied on non-parameteric bootstrapping methods to compare the means of distributions of fatalities [26, 27]. For events that do, and do not, involve weaponry banned during the FAWB, we drew 10,000 bootstrapped distributions, sampled with replacement from the original distributions. For each sampled distribution we calculated the mean,

and determined how often the mean of the distribution of events that did not involve such weaponry was greater than the mean of the distribution of events that did. This fraction forms the p-value for the one-sided comparison test [26, 27].

References

- [1] James Alan Fox and Emma E Fridel. The tenuous connections involving mass shootings, mental illness, and gun laws. *Violence and gender*, 3(1):14–19, 2016.
- [2] Louis Klarevas. *Rampage nation: Securing America from mass shootings*. Prometheus Books, 2016.
- [3] Lin Huff-Corzine, James C McCutcheon, Jay Corzine, John P Jarvis, Melissa J Tetzlaff-Bemiller, Mindy Weller, and Matt Landon. Shooting for accuracy: Comparing data sources on mass murder. *Homicide Studies*, 18(1):105–124, 2014.
- [4] Sherry Towers, Andres Gomez-Lievano, Maryam Khan, Anuj Mubayi, and Carlos Castillo-Chavez. Contagion in mass killings and school shootings. *PLoS one*, 10(7):e0117259, 2015.
- [5] James Alan Fox and Jack Levin. Mass confusion concerning mass murder. *The Criminologist*, 40(1):8–11, 2015.
- [6] Jason R Silva and Joel A Capellan. The media’s coverage of mass public shootings in america: fifty years of newsworthiness. *International Journal of Comparative and Applied Criminal Justice*, 43(1):77–97, 2019.
- [7] Grant Duwe. Body-count journalism: The presentation of mass murder in the news media. *Homicide Studies*, 4(4):364–399, 2000.
- [8] James Alan Fox, Jack Levin, and Emma E Fridel. *Extreme killing: Understanding serial and mass murder*. Sage Publications, 2018.
- [9] James Alan Fox and Jack Levin. Mass murder: An analysis of extreme violence. *Journal of Applied Psychoanalytic Studies*, 5(1):47–64, 2003.
- [10] Grant Duwe. Mass murder in the united states: A history. *McFarland & Company, Inc*, 2007.
- [11] William J Krouse and Daniel J Richardson. Mass murder with firearms: Incidents and victims, 1999-2013. Congressional Research Service, Library of Congress, 2015.
- [12] Raymond W Kelly. *Active Shooter: Recommendations and analysis for risk mitigation*. The New York City Police Department, 2012.
- [13] Adam Lankford. A comparative analysis of suicide terrorists and rampage, workplace, and school shooters in the united states from 1990 to 2010. *Homicide Studies*, 17(3):255–274, 2013.

- [14] Sherry Towers, Anuj Mubayi, and Carlos Castillo-Chavez. Detecting the contagion effect in mass killings; a constructive example of the statistical advantages of unbinned likelihood methods. *PloS one*, 13(5):e0196863, 2018.
- [15] D Wayne Osgood. Poisson-based regression analysis of aggregate crime rates. *Journal of quantitative criminology*, 16(1):21–43, 2000.
- [16] James O Lloyd-Smith. Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PloS one*, 2(2):e180, 2007.
- [17] Chester Ittner Bliss and Ronald A Fisher. Fitting the negative binomial distribution to biological data. *Biometrics*, 9(2):176–200, 1953.
- [18] Francis J Anscombe. Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika*, 37(3/4):358–382, 1950.
- [19] Glen Cowan. *Statistical data analysis*. Oxford university press, 1998.
- [20] Mark Gius. The effects of state and federal gun control laws on school shootings. *Applied economics letters*, 25(5):317–320, 2018.
- [21] Mark Gius. The impact of state and federal assault weapons bans on public mass shootings. *Applied Economics Letters*, 22(4):281–284, 2015.
- [22] RL Plackett. The truncated poisson distribution. *Biometrics*, 9(4):485–488, 1953.
- [23] MR Sampford. The truncated negative binomial distribution. *Biometrika*, 42(1/2):58–69, 1955.
- [24] AN Krishnan Nair. Distribution of students’*t*’and the correlation coefficient in samples from non-normal populations. *Sankhyā: The Indian Journal of Statistics*, pages 383–400, 1941.
- [25] Gregory W Corder and Dale I Foreman. *Nonparametric statistics: A step-by-step approach*. John Wiley & Sons, 2014.
- [26] Julie A Barber and Simon G Thompson. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Statistics in medicine*, 19(23):3219–3236, 2000.
- [27] Christopher Z Mooney, Robert D Duval, and Robert Duvall. *Bootstrapping: A non-parametric approach to statistical inference*. Number 94-95. Sage, 1993.