Mass Shootings in the USA

Hannah Reitz & Sophie Hensgen

2020-06-10

Contents

Introduction **1**

Theory and Hypotheses **2**

Data Gathering **4**

Data from Wikipedia 4

Data from GTrends 6

Data Analysis **7**

References **14**

lala

dfjhv Citation examples for in-text citations:

Smith cited [@smith04]. Smith cited without author [-@smith04]. @smith04 cited in line.

Introduction

Aufhänger: Many studies about media salience of mass shootings, but no research about how this media attention influences peoples' interest in certain events.

Based on previous definition of mass murder (citations from Duwe 2000, p.373) mass shootings in this paper are defined as "incidents in which four or more persons are murdered within a 24 hour period" [@Duwe2000, p. 323] with a gun.

Theory and Hypotheses

Kleine Einleitung in Literature Review:

“Like acts of terrorism and mass murder, not all mass shootings receive equitable media coverage, despite the sensational nature of the phenomenon as a whole (Schildkraut 2014, 2016). Absent from the body of research, however, is an exploration as to why this coverage is disparate” Schildkraut 2018 p. 224

Wie können wir einen Graphen einbinden/erstellen der unsere Zusammenhänge illustriert:

Deaths/Injuries/Race of perpetrator/ ——> The decline of the search interest Rich area/ Loose gun control in google trends is slower I /
I I I I V I more media attention for a longer time span due to more —> More people notice the newsworthiness of event event in media reports

More dead people –> more media attention –> higher attention in Google trends

newworthiness in den Medien

Könntest du evtl. gebrauchen: “Because agenda-setting theory is applicable at both the issue and attribute levels,17 media coverage of school shootings can be analyzed at these two levels as well. Although no particular theory was mentioned in it, a study examining the salience of fourteen school shooting incidents in network television news can be considered an example of issue agenda-setting research.18 By measuring the amount of news coverage of the incidents, the researchers showed that shooting incidents that resulted in more deaths or greater injuries were covered more. They also found that the amount of coverage was highly correlated across major television networks, demonstrating their “herd mentality.” Sung-Yeon Park (2012), p. 478

1. The more people got injured, the longer the shooting stayed in the trends.
2. The more deaths, the longer the shooting stayed in the trends.

evtl. newswothiness in Kombination mit inequality

Generally, research indicates that news coverage particularly is skewed toward stories about crime Schildkraut p. 223

The first two hypotheses mainly focus on how certain crimes in this case mass shootings garner more attention from the media because they are more “newsworthy” [@someauthors]. We argued that the more people died or got injured during a mass shooting the more newsworthy the event is and the more the media reports about it. In turn this high media attention leads to an increase in public interest and the decline of search hits in Google trends on this mass shooting is less steep than for mass shootings that are not as newsworthy, do not get reported on as heavily and do not accumulate as much public interest. The next three hypotheses concentrate more on specific characteristics of mass shootings and how those could affect the decline of the relative count of search hits on this event in Google.

First, the racial or ethnic identity of the perpetrator might be an important characteristic to consider as this could heavily shape the way media presents mass shootings and the ensuing public interest in the events. Mainstream media targets mostly the white middle class in th US as the goal is to reach a larch audience. Because of this most media outlets report from a white perspective and thus "non-whites are more likely to be portrayed in a negative light relative to whites" [@Callanan2012, p. 94] [@Tukachinsky2015, p. 187; @Wheeler2017, p. 73]. Therefore, a selection bias in media reporting in relation to race exists. Previous studies show that race explains some of the selection bias in the media. If the perpetrator in a mass shooting is not white or from another ethnicity than the majority population, the media covers the event more extensively [citations] Lundman

In line with these findings, we suggest that as non-white violators in mass shootings attract more attention from the media, this might lead to a higher interest in the event by the public. In light of the present study this higher interest in mass shootings committed by non-white violators by the broad public may manifest in a decline in hits on Google that is less steep than for events with white shooters. Hence, we expect that if the shooter is non-white or has a migration background, the decline in the relative count of search hits on the event is less steep. The third hypothesis, therefore, is:

H3: If the shooter is black or has a migration-background, the decline in the relative count of search hits on the event is less steep.

Another important characteristic to explain why the public interest declines slower for some mass shootings but not for others is the affluence of the area in which the shooting happens. There is not much research on the relationship between the Affluence of the area in which a mass shooting takes place and the media coverage it receives yet. However, the study by [-@Schildkraut2018, p. 235] shows that shootings that happen in more affluent communities receive more attention from newspapers. Accordingly, we assume that mass shootings in more affluent communities are more newsworthy and thus draw more media attention. Subsequently, the public interest in these shootings is greater and the reduction of the relative search hit count is less steep than for shootings that happen in less affluent communities. Therefore, the fourth hypothesis is:

H4: The more affluent a community a mass shooting happens in is, the shallower is the reduction of the relative search hit count.

As for the last hypothesis there is sparse research on how the gun laws of the state a mass shooting happens in influence its newsworthiness, media coverage and how this might affect the public interest. We expect that if a mass shooting happens in a state with weak gun laws the issue of gun control becomes more salient. Subsequently in light of this public discussion the event becomes more newsworthy and the media covers the shooting more heavily. Because of this, the public interest in the mass shooting rises and the relative count of the search hits on Google declines slower. Hence, the fifth hypothesis is:

H5: If the state a mass shooting happens in has weak gun laws, the relative count of the search hits on Google declines slower.

In summary, most of the previous research on mass shootings mainly focused on how certain characteristics of shootings affect its newsworthiness and its coverage in the media. Our contribution to this literature is to consider how certain characteristics of mass shootings could influence the public interest in form of the temporal decline in the relative count of search hits on this event on Google (bessere Formulierung?). Additionally we also consider characteristics of mass shootings previous studies have not considered. The focus of past research was mainly on the impact the count of dead and injured persons as well as the race of the perpetrator have on the media reporting. We also take into account how the affluence of the community where a shooting happens and the strength of the gun laws in the state a shooting takes place affect the public interest in the event.

Data Gathering

In order to analyze our hypotheses we combine data from Wikipedia that includes general information on mass shotings in the US with Google Trends data on the relative number of searches on Google for each event. A regression of the independent variables stated in the hypotheses on the decline of the relative search hit count on Google is conducted. This process is explained in detail in the following.

Data from Wikipedia

First, we scraped all tables from the Wikipedia page on mass shootings in the US which are available from 2000 to 2020 [@wiki] and cleaned the data from unnecessary information. Then the scraped tables were joined into one data set and all columns that include missings were dropped as we cannot include them in a regression. The resulting table includes information about the date, the location, the count of dead persons, the count of injured persons, the total number of dead and injured persons and a description for each massshooting.

Because we are using the data set based on the information from wikipedia later on to scrape data from google trends we created additional variables. We generated a variable with the dates fourteen days after the date of the mass shooting. This variable was subsequently used to create another variable which included both, the date of the mass shooting and the date fourteen days after the mass shooting in the required style of the Google Trends API. Furthermore, we generated a variable for the search term the Google Trends API is using to find information on the relative count of hits this keyword received in the 15 days time period. We conducted a manual search on Google Trends to find the search term for each event that received the highest number of hits in the 15 days time span for each event. Moreover, we discarded the years 2000, 2001, 2002 and 2003 as it is not possible to scrape Google Trends data before the year 2004.

After exporting this data table to excel we added three variables with information on the ethnicity of the shooter, the affluence of the area/community a shooting happened at and the strength of the gun laws for the state in which a shooting happened. These variable represent the indpendendent variables for the regression models in the analysis section additionally to the count of dead and injured persons. In order to find the necessary information on these characteristics, a online search was executed. The information on the ethnicity of the perpetrator was found on the respective Wikipedia page for each mass shooting. The variable was coded 0 if the shooter was black or had a migration background and 1 for white US Americans. The affluence of the community a mass shootings happened in was determined by a chart that depicts how the US Census Bureau breaks down income levels [@income]. ………………….. to be continued

Erlärung welche Daten, wie gescrapt, wie sehen die Daten aus Deskription Tabelle mit für Missing, welche observations haben wir warum gedropped, wie viele observations verbleiben in unserem sample –> timeperiods droppen.

```
url <- "https://en.wikipedia.org/wiki/List_of_mass_shootings_in_the_United_States"
t <- read_html(url)
table <- t %>% html_nodes("#mw-content-text > div > table") %>% html_table(trim=T)


table
```

hier Tabellen zusammensetzten

erstes klappt, danach gibt es ein problem mit incompatiblen types. Ich glaube wir müssen die buchstaben aus injured rausholen damit wir es in einen integer umändern können

To Do:

- gsub loop –> alle alle nicht zahlen rausziehen
- alle mutaten damit alle integer sein
- wenn das gemacht alle joinen zu einer großen tabelle
- in excel Extra Variablen hinzufügen
- Datum in Zahlen ändern
- zwei neue variablen, eine zwei Tage vor dem Event und 14 Tage danach
- Description den Namen herausnehmen

Erwähnen, dass wir nur nach einem Begriff suchen Search Terms verglichen bei Trends Alle Events, die über Tage gehen gedropped Sophie: 6-10 subben Hannah: 11-15

Delisle tripple murder droppen –» nicht genug daten Bart shooting evtl. rausnehmen –> sieht aus als hätte es sehr kleine Fallzahlen und einen seltsamen Verlauf Meteor shooting sieht komish aus –> evtl droppen wegen kleiner Fallzahl

Data from GTrends

Gtrends –> scraping von mehreren Namen auf einmal –> Text ob es mit einem funktioniert

```
#shootings <- gtrends(c("Milwaukee shooting", "Grantsville shooting", "El Paso shooting"), geo = "US",
#plot(shootings)
#shootings
```

Versuch eines Loop der einzelne outcomes speichert

```
#listshooting <- list()
#for (i in 1:nrow(wikishooting)){
#    x <- as.vector(wikishooting$SearchTerm[i])
#    y <- as.vector(wikishooting$TrendPeriod[i])
#    shooting <- gtrends(keyword = x, geo = "US", time = y, low_search_volume = T)
#    iteration <- (1 + length(listshooting))
#    listshooting[[iteration]] <- shooting
# }
```

```r
# Beispiel wie man die einzelnen Tabellen verbindet
#massshooting <- listshooting[[1]]$interest_over_time
#for (i in 2:length(listshooting)){
#    shoot <- listshooting[[i]]$interest_over_time
#    massshooting <- rbind(massshooting, shoot)
#}
#massshooting

#saveRDS(massshooting, "massshooting.rds")

# Load the city object as city
gtrends <- readRDS("massshooting.rds")
gtrends
```

First, we transform the `data.frame` into a `tibble`.

Data Analysis

Wie sollten vllt. Gender of shooter kontrollieren. Frauen werden in den Medien mehr beachtet –> Könnte zu einer Verzerrung der Ergebnisse führen.

```r
# Regression

regdead = lm(coef ~ Dead, data = Final, )
summary(regdead)
```

```
##
## Call:
## lm(formula = coef ~ Dead, data = Final)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4712 -0.6125  0.0695  0.4910  4.6038
```

```
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.842555   0.116000 -33.125   <2e-16 ***
## Dead        -0.007828   0.011079  -0.707    0.481
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9917 on 129 degrees of freedom
## Multiple R-squared:  0.003856,   Adjusted R-squared:  -0.003866
## F-statistic: 0.4993 on 1 and 129 DF,  p-value: 0.4811
```

```r
reginjured = lm(coef ~ Injured, data = Final, )
summary(reginjured)
```

```
##
## Call:
## lm(formula = coef ~ Injured, data = Final)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4581 -0.5623  0.0989  0.4983  4.6203
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.908371   0.089643 -43.599   <2e-16 ***
## Injured      0.001159   0.002322   0.499    0.619
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9927 on 129 degrees of freedom
## Multiple R-squared:  0.001927,   Adjusted R-squared:  -0.00581
## F-statistic: 0.2491 on 1 and 129 DF,  p-value: 0.6185
```

```r
regeth = lm(coef ~ Ethnicity, data = Final, )
summary(regeth)
```

```
##
## Call:
## lm(formula = coef ~ Ethnicity, data = Final)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3604 -0.6062  0.1468  0.4885  4.4730
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.7587     0.1318 -28.519   <2e-16 ***
## Ethnicity    -0.2416     0.1742  -1.387    0.168
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9863 on 129 degrees of freedom
## Multiple R-squared:  0.01469,    Adjusted R-squared:  0.007056
## F-statistic: 1.924 on 1 and 129 DF,  p-value: 0.1678
```

```r
reggun = lm(coef ~ Gunlaws, data = Final, )
summary(reggun)
```

```
##
## Call:
## lm(formula = coef ~ Gunlaws, data = Final)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4242 -0.5778  0.1008  0.4684  4.6508
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.8217     0.1479 -25.841   <2e-16 ***
## Gunlaws      -0.1148     0.1825  -0.629     0.53
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9921 on 129 degrees of freedom
## Multiple R-squared:  0.003059,   Adjusted R-squared:  -0.004669
## F-statistic: 0.3958 on 1 and 129 DF,  p-value: 0.5304
```

```r
regarea = lm(coef ~ Area, data = Final, )
summary(regarea)
```

```
##
## Call:
## lm(formula = coef ~ Area, data = Final)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5365 -0.6347  0.0385  0.5409  4.5385
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -4.0968     0.1667  -24.58   <2e-16 ***
## Area          0.2726     0.1947    1.40    0.164
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9862 on 129 degrees of freedom
## Multiple R-squared:  0.01497,    Adjusted R-squared:  0.007332
## F-statistic:  1.96 on 1 and 129 DF,  p-value: 0.1639
```

```r
regtotal = lm(coef ~ Total, data = Final, )
summary(regtotal)
```

```
##
## Call:
## lm(formula = coef ~ Total, data = Final)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4601 -0.5630  0.0971  0.4947  4.6167
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.9073765  0.0931309 -41.956   <2e-16 ***
## Total        0.0006171  0.0020209   0.305    0.761
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9933 on 129 degrees of freedom
## Multiple R-squared:  0.0007223,  Adjusted R-squared:  -0.007024
## F-statistic: 0.09325 on 1 and 129 DF,  p-value: 0.7606
```

```r
regfull = lm(coef ~ Dead + Injured + Ethnicity + Area + Gunlaws, data = Final, )
summary(regfull)
```

```
##
## Call:
## lm(formula = coef ~ Dead + Injured + Ethnicity + Area + Gunlaws,
##     data = Final)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3513 -0.5618  0.1065  0.4949  4.4237
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.714458   0.236470 -15.708   <2e-16 ***
```

```
## Dead        -0.023254   0.014677  -1.584   0.1156
## Injured      0.004641   0.003085   1.505   0.1349
## Ethnicity   -0.313963   0.175338  -1.791   0.0758 .
## Area         0.326111   0.195621   1.667   0.0980 .
## Gunlaws     -0.190793   0.182142  -1.047   0.2969
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.978 on 125 degrees of freedom
## Multiple R-squared:  0.06125,    Adjusted R-squared:  0.0237
## F-statistic: 1.631 on 5 and 125 DF,  p-value: 0.1565
```

```r
library(lattice)
library(MASS)
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select
```

```r
require(memisc)
```

```
## Loading required package: memisc

##
## Attaching package: 'memisc'

## The following object is masked from 'package:purrr':
##
##     %@%

## The following object is masked from 'package:tibble':
##
##     view

## The following object is masked from 'package:ggplot2':
##
```

```
##     syms
## The following object is masked from 'package:lubridate':
##
##     is.interval
## The following object is masked from 'package:rlist':
##
##     List
## The following objects are masked from 'package:dplyr':
##
##     collect, recode, rename, syms
## The following object is masked from 'package:rvest':
##
##     html
## The following object is masked from 'package:xml2':
##
##     write_html
## The following objects are masked from 'package:stats':
##
##     contr.sum, contr.treatment, contrasts
## The following object is masked from 'package:base':
##
##     as.array
```

```r
library(pander)
#regtable <- mtable('Model 1' = regdead,
#             'Model 2' = reginjured,
#             'Model 3' = regfull,
#             summary.stats = c('R-squared','F','p','N'))

pander(regfull)
```

Table 1: Fitting linear model: coef ~ Dead + Injured + Ethnicity + Area + Gunlaws

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-3.714	0.2365	-15.71	2.274e-31
Dead	-0.02325	0.01468	-1.584	0.1156
Injured	0.004641	0.003085	1.505	0.1349
Ethnicity	-0.314	0.1753	-1.791	0.07578
Area	0.3261	0.1956	1.667	0.09801
Gunlaws	-0.1908	0.1821	-1.047	0.2969

```
#regdead <- lm(coef ~ Dead, data = Final, )
#report_regression(regdead)
```

References