Black Swan Shootings: A Model for Predicting the Worst of the Worst Mass shootings

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Matthew D'Anna
Master of Advanced Study
Arizona State University, 2008
Master of Arts
John Jay College of Criminal Justice, 2008

Director: Christopher S. Koper, Professor
Department of Criminology, Law, and Society

Summer Semester 2020
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

I would like to thank the family, friends, colleagues, and supporters who have made this happen. My loving wife, Wendy, helped keep me sane throughout the process. Dr. Koper, and the other members of my committee were of invaluable help. Madlib and J Dilla provided the soundtrack to make this possible. Now...my kids get their dad back!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ABSTRACT

BLACK SWAN SHOOTINGS: A MODEL FOR PREDICTING THE WORST OF THE WORST MASS SHOOTINGS

Matthew D'Anna, Ph.D.

George Mason University, 2020

Dissertation Director: Dr. Christopher S. Koper

Since the late 1990s, mass shootings in the United States, and in particular public mass shootings with high victim counts, have increased in frequency at an alarming and largely unprecedented rate. Most academic studies on mass shootings are descriptive and offender-centric, offering few tangible opportunities for predicting future mass violence. This study takes a different approach. This study develops a county-level spatial threat assessment for identifying locations at high-risk for experiencing or producing 'Black Swan shootings', defined here as an attack involving a perpetrator(s) using firearms to kill or injure a significantly large number of innocent or unwitting people, chosen intentionally or at random. Counties are evaluated for their risk of experiencing these attacks or producing attack perpetrators based on community-level measures for social contagion, public safety, demographics, mental health and substance abuse, and weapons availability. Using data from 18 different sources on mass shootings, 44 events since

1998 are identified. Based on statistical anomaly detection for extreme casualty counts across all mass shootings since 2013, Black Swan shootings are events with either eight killed, 13 wounded, or 15 total casualties. The analysis of these events indicates clear contagion effects over space and time: attacks occur in 17 distinct clusters of counties and most of the time occur within one year of the prior attack, with heightened risk in the 35 days immediately following an attack. Using annual county-level data from 1998 to 2018, results from t tests, Cohen's D effect sizes, and Mann-Whitney U tests indicate that communities which experience Black Swan Shootings or produce their perpetrators have statistically significant higher levels of violence, denser populations, more racial diversity, higher percentages of females, and higher counts of firearm laws compared to areas without such attacks. The spatial threat assessment uses a logistic regression model based on these most significant social factors. Retroactively, each year this model identifies less than 1% of the country as 'high-risk.' The model is deemed analytically viable when the metric score for the accuracy and precision of the count and location of Black Swan shootings in a given year is greater than 50%; this occurs in ten different years, all occurring since 2006. This includes nine specific bullseyes, where the model was an exact match for the county and year of either an attack or the offender residence of a Black Swan Shooting. The model's performance is evaluated, with a discussion on the opportunities for operationalizing these findings to inform on future attacks.

INTRODUCTION

Background

Since the terrorist attacks of September 11th, 2001, the United States has not had another foreign attack of the same magnitude and lethality. Instead, terror has a new identity, as the U.S. has since experienced some of the most horrific, homegrown attacks in its history. Rather than another coordinated event like 9/11 leaving thousands killed across multiple attack sites, there has been an increase in singular homegrown attacks, often committed by “lone wolf offenders.” These attacks defy conventional predictability, motivation, and attribution. In the aftermath of 9/11, there has been a disturbing pattern of more attacks with increased casualty counts over shorter time periods, including:

- On 4/16/2007, a single attacker armed with two handguns killed 32 people and wounded 23, on the campus of Virginia Tech University in Blacksburg, Virginia.
- On 7/20/2012, a single attacker armed with two handguns, an assault rifle, and a shotgun killed 12 people and wounded 70, at a movie theatre in Aurora, Colorado.
- On 12/14/2012, a single attacker armed with two handguns, an assault rifle, and a shotgun killed 28 people and wounded two at Sandy Hook Elementary School in Newtown, Connecticut.
- On 6/12/2016, a single attacker armed with a handgun and a rifle killed 50 people and wounded 53 at Pulse nightclub in Orlando, Florida.

- On 10/1/2017, a single attacker armed with at least 23 firearms killed 59 people and wounded 441 at an outdoor concert in Las Vegas, Nevada.

Each of these is considered a mass shooting. A particularly American brand of killing, mass shootings are loosely defined as the killing or injuring of four or more people with a firearm, and are at an all-time high in the U.S. While debates on definition fuel arguments of whether activity is truly increasing (Fox & Levin, 2015; Duwe, 2004), the increased media and political attention has brought the most extreme mass shootings into the mainstream, national spotlight (Duwe, 2007). Some contend that the rate of occurrence, which by some definitions is daily, is reaching a volatility worthy of a social pandemic (Krouse & Richardson, 2015). As Goode and Ben-Yehuda (1994) describe, a moral panic need not be a rational response, but intense feelings driven by the perception of harm. To address these perceptions, with each new attack there is a charged debate on proposed legislative changes related to firearm acquisition and ownership. Meanwhile, gaps in contemporary research limit awareness on the causes and patterns related to these events, as well as general firearm violence as a public health issue.

This study contends that there are two types of mass shootings: 'ordinary' mass shootings, occurring relatively frequently and involving casualty counts on the lower end of the definitional spectrum; and 'Black Swans,' the rarer, seemingly unpredictable attacks, often in public spaces and involving significantly higher casualty counts. The so-called ordinary mass shootings are no doubt tragic and warrant proper attention and research; however, those events are not the focus here. The notion of a Black Swan, as originally described by Taleb (2010), is an event that is seemingly impossible to predict and has catastrophic consequences. These are the outliers - high-profile, rare events that come as a surprise - in fields such as history, science, and finance (Taleb, 2010). In this context, Lankford and Silver (2019) demonstrate that since the 1960s public mass shootings have increased in count and lethality. In fact, when examining such attacks based on thresholds of 8, 12, or 16 victims killed, Lankford and Silver (2019) found that 2010-2019 accounts for more than 50% of the deadliest public mass shootings. While still rare (no decade has more than 18 attacks, and prior to the 2010s no decade had more than five attacks) and warranting the 'black swan' label, Lankford and Silver (2019) set the foundation that public mass shootings are a growing threat in the U.S.

Most prior analysis and research on the extreme mass shooting events focuses on the individual offender, not on the community surrounding the attack. Currently there exists a lack of evidence related to pattern identification, commonalities, and predictive power for understanding these attacks. This study aims to identify the common social factors of communities experiencing Black Swan Shootings and use those patterns to assess the communities at highest risk for future events to occur. The end result is designed to be a proactive model that determines the location-based layers with the strongest relationships to Black Swan Shootings and is able to identify the communities for future extreme attacks to occur. If successful, this model could serve as a tipping and queuing mechanism for enhanced community engagement across social organizations, legislation, and law enforcement, among others. If this study identifies the communities, additional research could leverage micro-geographic analyses to narrow the specific target locations and possible offenders within such communities.

The social factors to be examined are measures related to violence, socioeconomics, mental health, and firearm preferences. This analysis seeks to identify whether combinations of the layers derived from each of these four variables are able to predict the highest risk areas for a future Black Swan Shooting, and if so, to what extent. This method leverages the concepts of environmental backcloths of crime, where events are driven by the macro-level location-based correlates of features to determine the spatial vulnerability of an area (Brantingham and Brantingham, 1993). The resultant landscape identifies potential new target areas, based on their similarity to known risk locations. Such a process involves identifying the relationship of each geographic feature to the events, and then determining the predictability of any combination of factors, through a multivariate regression model. The findings, if significant, can be applied to the current environment and time period, to identify "high-risk" communities.

The underlying assumption of this study is that despite assessments and conjecture to the contrary - yes, patterns do exist among the deadliest contemporary mass shootings. These patterns are not clear-cut and obvious, but they do exist. To find these patterns and correlates, typical behavioral measures such as intent, motivation, and affiliation are intentionally excluded from this analysis. As will be described in the literature review, multiple studies have addressed the lack of predictability for such attacks using offender traits and behaviors. Thus, the focus here is on the communal and societal factors, and their influence on attack occurrence.

Thus, there are three hypotheses for this study. First, this study contends that degrees of commonalities exist among community characteristics related to violence, socioeconomics, mental health, and firearm preferences for communities experiencing and likely to experience Black Swan Shootings. Second, while the strength of these relationships are expected to vary, the combined layering of statistically significant unique features of these communities is expected to highlight a pattern that can be used to identify the communities at highest risk for future attacks. Third, the combination of these commonalities is different from those communities without attacks; thus, commonalities among communities not experiencing, and less likely to experience, such attacks will also exist. Therefore, the most significant commonalities should be able to identify a subset of communities with an elevated risk for a future Black Swan Shooting.

Definitions

Mass Violence

Defining Black Swan Shootings involves components of mass murder and mass shootings. Unfortunately, across academic research there has been a lack of consistency when it comes to defining these terms. Due to these inconsistencies, it is difficult to pinpoint an exact definition as to what constitutes and separates mass murder from spree murder and serial murder (Palermo, 1997). According to Busch and Cavanaugh (1986), there are subsets of multiple murder: mass murder is multiple murders in a short time frame (minutes to hours); spree killings are multiple murders connected to one event over a time period of hours to days without a break/cooling off period, which may include mass and serial murders; serial murder is multiple murders with cooling off periods, of two days or more; familicide is multiple family member murders in a short time period;

and sex-related murder is sexually motivated killing, determined via evidence. Palermo and Ross (1999) found intent is what separates mass murderers from spree and serial murderers. It is worth noting the overlap in these definitions and typologies. Thus, timing may be viewed as a key characteristic when distinguishing serial, spree, and mass murder (Fox and Levin 2003), where "cooling off" helps differentiate the murder types.

The Federal Bureau of Investigation (F.B.I.) defines mass murder as events resulting in at least four deaths and normally taking place at one or more geographic locations relatively near one another (F.B.I., 2018). Similarly, Dietz (1986) and Fox and Levin (1998) stress a 24-hour or less constraint, along with a reasonable distance between attacks. When defining mass murderers, several studies limit events to a single offender (Gresswell & Hollin, 1994; Meloy et al., 2001; Hempel & Richards, 1999). Lankford (2015b) found offenders are more often concerned with inflicting harm than protecting themselves, and similarly Meloy et al. (2004) describe indiscriminate killing of random victims. However, Fox and Levin (2003) note the more common mass killings involve known victims in non-public places. Mass murderers are typically associated with firearms, but Meloy et al. (2001) and Fox and Levin (2015) found 33% of events involve weapons such as fire, explosives, knives, and blunt objects. Dillon (2013) notes other weapons can include bombs, poison, and even choking.

Regarding the dynamics of the attack, mass murders can have varying degrees of spacetime windows. Multiple studies advocate for one place at one time (Petee et al., 1997; Palermo & Ross, 1999; Ho Shon & Roberts, 2010; Knoll, 2010; Knoll, 2012; Kennedy-Kollar & Charles, 2013), and others describe it as a "single incident" (Meloy et al., 2004; Hempel & Richards, 1999; Palermo, 1997; Lankford, 2015; Holmes & Holmes, 1992; Fox & Levin, 2003; Lankford 2015b; Fox & DeLateur, 2014). Other studies allow for bifurcated attacks at multiple locations (Aitken et al., 2008; Huff-Corzine et al., 2014; Meloy et al., 2004). In general, the attack timeline for a mass murder is considerably shorter than a serial murderer. The F.B.I. treats the spacetime concepts of "one event" and "close geographical proximity" as having no cooling off period between murders. Such a cooling off, or significant temporal break, is what differentiates public mass killing as a spree killing instead of a serial killing (Lemieux, 2014; Chapman, et al. 2006; Aitken et al., 2008). Some studies consider this to be a few minutes to a few hours (Fox & Levin, 1998; Carcach et al., 2002; Fox & Levin, 2003; Huff-Corzine et al., 2014), and others describe it as a 24-hour window (Hilal et al., 2014; Duwe, 2000; Bowers et al., 2010; Duwe, 2005; Allely, et al., 2014; McPhedran & Baker, 2011). Overall, an attack is expected to last for less than a day.

Further, consider what mass murderers are not. According to Duwe (2004), mass murderers are not involved in riots, lynching, or other collective violence. Fox and Levin (2003) rule out war crimes, large-scale political terrorism, and some organized crime activity. Similarly, Levin and Madfis (2009) have defined mass murder as being antisocial and non-state sponsored. In this study, genocide and other state-sponsored forms of violence as separate from mass murder, because as Liwerant (2007) notes, political systems and their associated structural and systematic destruction are considered the root of those crimes, not the individual. Mass murder excludes other types of homicide, including serial murder, spree murder, felony-related, gang motivated, politically motivated, war crimes, organized crime, political terrorism, riots, and lynching.

Ultimately, Holmes and Holmes (1992) argue that a strong mass murder definition should consider the victim count, the locations, the timing of the killings, and any distance between attack sites. Those variables will determine spree versus serial versus mass murder. Mass murder can be described as the sudden and catastrophic indiscriminate killing a group of people in a single event of one or more closely related locations, typically within 24 hours or less.

Mass shootings are not distinctly different than mass murder. In general, mass shootings are a subset of mass murder, where the circumstances surrounding the attack are a mass murder with a firearm. Over time, mass shootings have become the most popular form of mass murder (Duwe, 2004). All Black Swan Shootings are mass shootings. Most definitions use a four-victim count, but vary on whether those victims are fatalities, or counts of fatalities and wounded (Krouse and Richardson, 2015).

Mass public shootings are a type of mass shooting, and thus another sub-type of mass murder, that excludes residential familicides. The F.B.I. uses the same descriptors for mass murder as mass public shootings, just with an added weapon type (Lemieux, 2014). Many, but not all, Black Swan Shootings are expected to be mass public shootings. Some studies exclude shootings that take place during the commission of other crimes, such as robbery, gang violence, hostage-taking, criminal competition, insurance fraud, arguments, romantic triangles, and domestic violence (Follman et al., 2018; Krause and Richardson, 2015; Cannon, 2016; Lankford, 2016; Lankford, 2015b). Mass public shootings get a significant amount of media attention, and have seemingly the highest potential for devastation, but are relatively rare occurrences. Part of the media attention, and subsequent moral panic discussion, involves the seemingly indiscriminate attack on innocent bystanders (Lankford, 2016; Lankford, 2016b; Dahmen et al., 2017; Bjelopera et al., 2013). Attack location aside, these events are thought to have different motivations and methods than other mass killings (Dillon, 2013; Lankford, 2016b). For example, a rampage school shooting is a type of mass public shooting that occurs at a school involving current or former attendees, where victims can be either symbolically targeted or shot at random (Newman and Fox, 2009). Mass public shootings can be at single or multiple locations (Krouse & Richardson, 2015). Common attack sites include schools, workplaces, restaurants, places of worship, parking lots, malls, shopping centers, public transit, and even private parties (Bjelopera et al., 2013; Krouse & Richardson, 2015; Lankford, 2016). Mass public shooters are thought to have done extensive planning and are willing to kill symbolic targets. One of the best ways to understand the widespread panic associated with mass public shootings, according to Lankford (2016b), is to treat them as functionally similar to terrorism.

Aside from event classification, there are two casualty-related components to consider when defining events for this study: first is casualty counting, second is whether to include the perpetrator(s) in the casualty counts. Regarding counting casualties, previous mass murder and mass shooting research is relatively mixed. Multiple recent studies rely exclusively on fatality counts for examining mass murders and mass shootings (Hilal et al., 2014; Dahmen et al., 2017; Cohen et al., 2014; Team Trace, 2017;

Fox & Levin, 2015; Fox, 2013; Everytown, 2018; Krouse & Richardson, 2015; Lankford, 2016; McPhedran & Baker, 2011; Dillon, 2013; Cannon, 2016; Follman et al., 2018; F.B.I., 2018; Lankford, 2016b; Chapman et al., 2006). Some of these studies use fatalities-only to stay consistent with the F.B.I. (Lankford, 2016b); some do so as a derivative on the dataset in their analysis (Dillon, 2013); and others attempt to stay consistent with historical studies that did not have access to accurate wounded victim counts (Fox & Levin, 2015; Fox, 2013). Other recent literature examines mass murders and mass shootings as counts of fatalities and wounded victims (Kleck, 2016; Holmes & Holmes, 1992; Dietz, 1986; Lankford, 2015b; Lowe & Galea, 2015; Bjelopera et al., 2013). From a shot fired perspective, Kleck (2016) finds fatalities versus wounded to be relatively meaningless, stressing that what matters is how many rounds were fired. Kleck (2016) found that excluding wounded victims via F.B.I. Supplemental Homicide Reports led to missed cases of high wounded, low fatality counts. This study measures Black Swan Shootings as killed and wounded victims for several reasons. First, there is an inherent luck often associated with dying versus surviving an event, and it is inconsistent to measure the unlucky. Second, as technology platforms and digital media outlets have advanced, data quality has improved. Thus, injury counts appear to have improved in the modern age of communication, benefiting from streaming internet and 24/7 news cycles.

Separate from, but related to, casualty counts is whether to include the perpetrator as a victim. A majority of recent studies, using some of the most widely available and reliable official and open source datasets, exclude the shooter in victim counts (Meloy et al., 2001; Hempel & Richards, 1999; Huff-Corzine et al., 2014; Dahmen et al., 2017;

Fox, 2013; Everytown, 2018; Bjelopera et al., 2013; Chapman et al., 2006; Follman et al., 2018). The counterargument, to include the shooter as a victim, is to take into account the mental health of the perpetrator and recognize their victimization. While obviously not condoning their actions, this perspective attempts to appreciate the suffering of the offender, that ultimately led to the attack occurring. For this study, the shooter(s) will be excluded in any casualty count calculations. As previously stated, this study avoids analysis of offender motivations and intent, and attempting to infer the intentionality of the perpetrator(s) desire to live or die is nearly impossible. There is too much inference, and not enough known details, to reasonably account for it in this study.

Black Swan Shootings

Black Swan Shootings borrow from key concepts of mass murder, mass shootings, public mass shootings, spree killings, and the traditional definition of a Black Swan event. A Black Swan Shooting is defined as an event occurring in the United States involving one or more coordinated perpetrators acting on behalf of themselves or a non-state-sponsored entity using one or more firearms to kill or injure a significantly large number of people.

A Black Swan Shooting is defined as a single streaming action of relatively short duration, that can encompass one or more attack sites over the course of minutes or hours. Although an event is treated as a fluid situation, it should be less than 24 hours and not extend beyond a single period of darkness in the locale. There should not be a distinct interval (or "cooling off") during the course of the event. An attack may occur in any combination of public and private spaces. The number of attackers does not matter;

however, they must be acting together and on behalf of themselves or a non-government group. Black Swan Shootings do not include state-sponsored terrorism or genocides, but may include non-state-sponsored terrorism, spree and serial murderers, other murders related to felony activity, politics, gangs, cults, and hostage situations. Black Swan Shootings are not fights, clashes, or other mutual attacks involving multiple sets of opposing attackers battling each other. For example, rival gangs shooting each other at a public venue does not qualify as an event for this study. Terrorism is included if the event involves a firearm, as well as potentially other weapons, and a high casualty count. If a terror attack involves a bomb alone, then it is excluded from this study. For example, the terrorist attacks of September 11, 2001 are excluded, as a firearm was not involved. Since motive of attack is irrelevant for case inclusion, homegrown and foreign-inspired terrorist attacks are included when the event matches the weapon and casualty count threshold. The weapons used must consist of, but are not limited to, a firearm. This involves any type of firearm, including handguns, assault rifles, shotguns, altered/manufactured firearms, and weapons using large capacity magazines. Other weapons, including improvised explosive devices or military grade munitions, may be included if used in conjunction with a firearm. Victim selection may be intentional or random, as the victim-offender relationships prior to the event are not significant in this construct. Thus, this study potentially includes familicides. As this definition attempts to identify mass carnage, regardless of the perpetrator's intent, casualty counts for an event are defined as the total number of fatalities and wounded/injured survivors, excluding the shooter(s). Ultimately, there can be a narrow distinction between wounded and killed. While injuries can have significant variance for severity, total attribution is both possible and the most comprehensive way to quantitatively score/assess/understand the full effect of an attack.

This definition is a combination of the aforementioned conventional definitions for mass casualty attacks, as well as accounting for the increased occurrence and severity unique to public mass shootings noted by Lankford and Silver (2019). What further distinguishes these types of events as “Black Swans” is their seemingly unpredictable nature (especially when considering failed offender profiling strategies), and exceptionally high casualty counts. The final components of the definition - the victim count threshold and temporal start - will be addressed in the methodology section.

Community-level social factors

While it is a reasonable assertion that mass shootings have increased over the past decade, the reasons why remain unknown. This study contends that location-based social factors may contribute to the causality. A proper location-based predictive model for Black Swan Shootings should involve the spatial influence of the most significant environmental features. To identify those features, concepts of vulnerability and exposure are measured from a macro-geographic perspective. Vulnerable areas are based on the presence of criminogenic features, and exposure is the level of such vulnerabilities in a community. Thus, in this context the landscape is the United States, and the threat is Black Swan Shootings. Currently, there are neither universal risk factors for mass shootings, nor for extraordinary events such as Black Swan Shootings. Unlike local crime at the micro-geographic level, where generations of partnerships between law enforcement, the community, and other government agencies have combined to better comprehend the spatial influences on crime, the spatial factors in Black Swan Shootings are unknown. Within the landscape and threat environment of this study, defining the key characteristics of a community consists of four variables: safety, demographics, stability, and weapons.

Safety within a community is defined by the perceived degree of risk and vulnerability an individual may associate with the local environment. This is best measured through levels of violence, and will be operationalized as violent crime, and law enforcement officer line of duty firearm fatalities. The demographics of a community are the characteristics that define the diversity of people; the characteristics that make neighborhoods similar and different. Demographics can be measured through socioeconomic variables, and in this study will be operationalized as population density, race, sex, poverty, unemployment, divorce, and head of household status. Third, and related to demographics, is stability. Stability is the relative strength of a community, measured by the severity of mental health conditions and abusive tendencies/behaviors. These are operationalized as suicides, and drug and alcohol overdoses. Fourth, a location-based perspective on firearms should address how a community feels about weapons. This is best examined through weapon availability and political sentiment. These measures will be operationalized as firearm ownership and firearm legislation. Thus, this study attempts to measure community levels of safety, demographics, stability, and weapons through changes in violent crime, socioeconomic factors, mental health indicators, and firearm preferences. Each of these components will be addressed in the literature review and methodology for their connections to relevant criminological theories, mass shooting research, and studies on location correlations with crime.

Prediction

Before proceeding to the literature review, a few thoughts on prediction. First, prediction is hard. At its core, prediction is making a guess about the future (Gottfredson & Gottfredson, 1988). Prediction is something we do on a daily basis - anticipating future behavior, looking for warning signs, and reading cues to determine what is going to happen next (Chaiken et al., 1994). When a prediction works, the appropriate combination of data and analysis has occurred. When it fails, further calibration is needed: more/different data, refined methods, or new questions. Within criminology, prediction has typically been critiqued for poor accuracy, lack of standards, and unreasonable goals (Gottfredson & Gottfredson, 1988; Chaiken et al., 1994). This may be because criminology prediction is often thought of as person-centric, identifying the individual causes of violent behavior, or assessing an offender's likelihood of recidivism. Such predictions are exceedingly complex and difficult to model, requiring data and resources far beyond criminology alone.

These criticisms do not mean criminological predictions should cease; rather, it means predictions need to improve the data, refine the methods, and reframe the questions. Further, causality and prediction are often erroneously linked; good predictions do not require causality, and predictions do not have to prove causation. Event-centric and place-based modeling versus person-centric guesstimation offers better opportunities to improve criminological predictions. Controlling for the offender and focusing on the environment allows for a better understanding, and more functional prediction, of criminal behavior.

Second, predicting rare events is even harder. Most often, on any given day the event count is zero - and there is potential for a large number of false positives. It can be easy to label rare occurrences as “unpredictable,” because of a low frequency of events or lack of available data. While some of this is absolutely true (there is no profile for mass shooters, as Bjelopera et al. (2013) correctly argues), some of this also stems from a lack of ingenuity. Yes, applying the same standard statistical measures to 30,000 burglaries as 30 homicides will certainly yield questionable results. However, a deeper investigation into the 30 homicides can create a richer dataset that identifies more attributes, more relationships, and more potential hidden subconscious decision-making processes among a set of heterogeneous offenders. At that point, the predictability for the 30 events has the potential to be much better than the 30,000 events, as the former set lacks the volume of the latter, but gains depth. As Paulsen et al. (2010) and Harding et al. (2012) describe, the world does not behave in a completely random way; patterns exist, and not always in the obvious datasets. Further, King and Zeng (2001) note that data collection matters. Efficient sampling of non-events helps balance the number of observations with properly calibrated explanatory variables.

Thus, identifying qualitative patterns among events, and enriching a low-count event dataset with macro-level factors, is more ideal. Such a process, which is not suitable for traditional criminological methods, is ideal for determining the potential causality of macro-level factors on events. Specific to this study, in examining mass murderers Dietz (1986) and Fox and Delateur (2014) noted the weakness in prediction models that focus on the offender and their personality and demographic traits. As previously mentioned, those are poor predictive values, because the offender often has a set of characteristics similar to a large portion of the general population. A dataset that focuses on event and geographic characteristics, may be able to properly inform and predict better.

Third, there is a difference between event prediction and risk assessment. Event prediction is fatalistic; attempting to determine the specific likelihood of occurrence. In such a model, predicting rare events are consistently zero, or null. However, risk assessment controls for behaviors, and scores for relative indicators. Thus, if all targets have a low probability, risk assessment still demonstrates the potential for an event, and identifies the key indications and warnings. Since the goal is prevention, not prosecution (Follman, 2015), a risk assessment can judge all targets to have a low probability of attack, but still rank targets on relative scales of magnitude. As Follman (2015) found, mass murder is actually better suited for risk and threat assessment; it is not considered an impulsive crime, meaning there are clear, measurable indicators (Follman, 2015). Unlike traditional offender predictions, these are not demographic or biographic traits, but consist of actions and events that slowly change. As Mulvey and Cauffman (2001) describe, these are not easy to observe, as they can require constant and dynamic monitoring and updating, but nonetheless possible. This notion of a pathway of indicators has been discussed for decades in criminological research (Gottfredson and Gottfredson, 1988), but rarely applied proactively.

Despite these concerns surrounding prediction, this study plans to do it. This study offers an opportunity to measure the social conditions of such pathway indicators and build on the recent work measuring area-level factors associated with mass violence and public shootings from Markowiak et al. (2018), Lowe and Galea (2015), and Kalesan et al. (2016). This design avoids offender-centric, biographic predictions; rather, this study is designed to model location-based factors of Black Swan Shootings and use those relationships to assess community-level risk for attacks and residences. Academic research is often not suited for real-time, persistent updating. This study hopes to change that paradigm and offer an actionable threat assessment on Black Swan Shootings that focuses on macro-geographic characteristics rather than the offender traits.

LITERATURE REVIEW

A Black Swan Shooting is defined as an attack in the U.S. involving a perpetrator(s) acting in a non-state-sponsored capacity using a firearm(s) to kill or injure a significantly large number of people. As such, an analysis of prior academic studies related to Black Swan Shootings cuts across multiple forms of firearm violence and mass violence. A complicated relationship exists among firearm deaths, suicides, mass violence events, school shootings, and active shooter situations, where each type of attack has unique degrees of intersection with the others. There exists neither perfect inclusion or exclusion among the different types, and it is the non-overlapping elements of each that create the foundation for this study. For example, most analysis and research on mass shootings focuses on the individual offender and their behavioral traits, not on the geographic and community elements that this study is interested in. Most research on firearm deaths details such community and socioeconomic conditions but does not concern itself with specific offender traits. In the context of Black Swan Shootings, combining these unique research perspectives offers insight into the motivations, characteristics, and commonalities of such attacks across behavior, space, and time. As such, it would be short-sighted to not include previous research across each of these types. Thus, Figure 1 is an illustrative diagram that highlights the degrees of intersection among these research areas of focus relevant to this study.

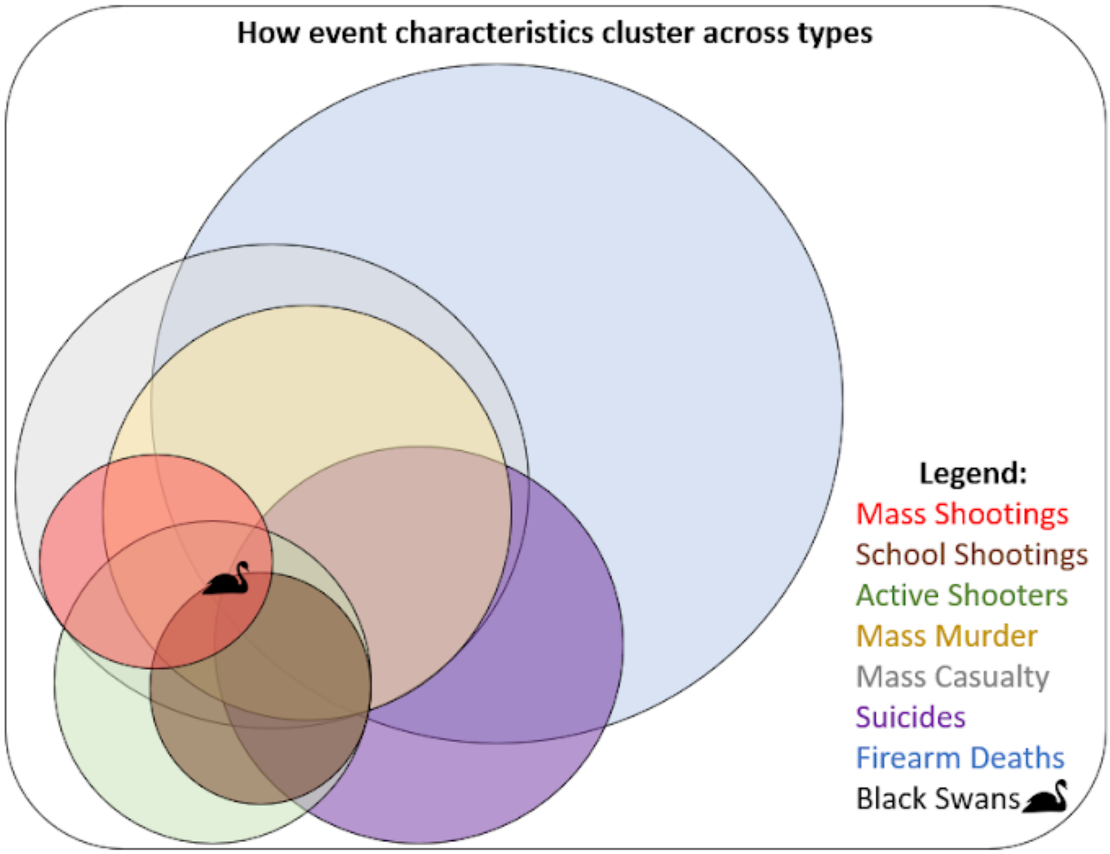


Figure 1 Murder research study illustration, clustered by type and scaled by volume of events

With this context, this literature review has three parts. First, it assesses prior research on events related to Black Swan Shootings, focusing on the prevalence and trends of similar attacks. This includes the attack frequency, offender traits and behaviors, and attack logistics. A theme in this section is the similarity of traits of such attackers to a large percentage of the general population. Such similarity makes those characteristics not distinctly unique or targetable for identifying potential future offenders.

Second, this literature review examines research on individual and personal causes of mass violence attacks. Despite having a location-based, macro-geographic approach to this study, this section addresses the non-geographic correlates and causality for mass violence by examining the perpetrator gender and mental health patterns.

Third, this literature review addresses societal and community causes of mass violence. This section includes general social factor trends, including media and exposure to violence, and copycat and contagion effects of mass violence. This section also includes state and community correlates to violence; firearm availability research; and studies on the effects of firearm legislation on firearm violence. Research on macro-geographic location-based factors examine the effect of violence on communities, socioeconomic correlates with violence, social disorganization, and public opinion that shapes attitudes and beliefs about firearms and violence. These place-based and geographic proximity studies demonstrate how location can typically be a better predictor than offender traits. Next, firearm availability informs our understanding of mass violence, as a majority of such attacks (and all Black Swan Shootings) involve firearms, and access to weapons is not uniform across populations and communities in the U.S. These studies demonstrate that increased access to weapons has the potential to increase the likelihood of a mass shooting. Regarding firearm legislation, this literature review covers the impacts of laws on general firearm violence and mass shootings, examining the impact of laws related to firearm use (right to carry, shall issue, stand your ground, and conceal carry), sales (seller/buyer restrictions, permit to purchase, background checks, and waiting periods), and possession (removals, bans, and restrictions). Firearm legislation is treated as a proxy for local firearm preferences and political sentiment. Similar to firearm availability, this research demonstrates that local conditions may impact the ability for a mass shooting to occur.

Mass violence prevalence, patterns, and trends

This section examines research on the prevalence and trends across types of mass violence, including mass casualty and mass murder, mass shootings, school shootings, and active shooters. The research from each attack type offers insights into the unique aspects of Black Swan Shootings. The prevalence and trends across attack types refers to event frequency, offender demographic traits, and attack components, which includes site selection and weapon use.

Event Frequency

Overall, mass violence in the U.S. has increased over time. The occurrence rates of mass murders in the U.S. from the 1970s to the mid 1990s averaged approximately two attacks per month (Fox and Levin, 1998). But as Lester (2002) found, that average can be misleading; since the 1970s single-victim murder has declined, while two-victim murders have increased. More recently, Krouse and Richardson (2015) found during 1999-2013 the U.S. experienced 31 mass murders annually, which is 2.6 attacks per month. Regarding mass shootings, using F.B.I. Supplementary Homicide Reports from 1999-2013, Krouse and Richardson (2015) identified an average of 21 attacks per year. Globally, during 2002-2016 mass shootings accounted for 10% of terrorist attacks yet caused more than half of the fatalities, with the U.S. leading the world in the proportion of terrorist attacks involving firearms (Tessler et al., 2017). Mass shootings are largely a U.S. problem, as Lankford (2015) found during 1966-2012, 31% of global public mass shootings occur in the U.S. As Krouse and Richardson (2015) note, the days between public mass shootings have steadily decreased by decade. In the 1970s, events occurred an average of 282 days apart; in the 1980s, 152 days; in the 1990s, 94 days; in the 2000s, 88 days; and from 2010-2013, 74 days apart. Follman et al. (2018) found since 1982 mass shootings have tripled in frequency, as 63% of attacks have occurred since 2006. Since the 1970s, Dillon (2013) and Krouse and Richardson (2015) found large increases in annual public mass shooting rates, rising from one or less per year to upwards for 3-4 attacks annually. Krouse and Richardson (2015) found in the past 50 years there have been 13 public mass shootings with ten or more fatalities, and seven of these 13 attacks have occurred in the 2000s, resulting in more than 50% of the casualties of all 13 events. With this rise in prevalence, 2009-2013 was considered at the time to be the deadliest time period for public mass shootings (Krouse and Richardson, 2015). More recently Lankford and Silver (2019) demonstrate that the years 2010-2019 has continued this trend, where multiple victim thresholds demonstrate that this time period has accounted for more than half of the deadliest mass shootings since the 1960s.

While overall levels of violence in schools are decreasing (O'Toole, 2000), and the risk for homicide is relatively low (Nekvasil and Cornell, 2015), school shootings have also increased since the 1990s (Rocque, 2012). Nekvasil and Cornell (2015) found 0.3% of homicides and 0.8% of multiple casualty homicides occur in schools. However, during 1974-2002 Newman (2007) identified 21 "rampage" school shootings - attacks at

U.S. high schools involving multiple homicides and random victims - which is nearly one per year. Similar to mass shootings and school shootings, Blair and Schweit (2014) have observed an uptick in recent active shooter attacks: during 2000-2006 there were 6.4 incidents annually, during 2007-2013 there were 16.4 active shootings annually, and more recently, there have been 50 active shooter events across 21 states during 2016-2017 (F.B.I., 2018).

However, patterns in mass violence attack frequencies are not linear. Leading up to the 2000s, Duwe (2007) identified distinct waves of mass murder trends: Wave 1, during 1900-1939; a gap during 1940-1965; and Wave 2, during 1966-1999. Quantitatively, Wave 1 and Wave 2 are similar - mass murder occurred as often during 1920-1940 as it did into the 1980s. Qualitatively, Wave 1 almost exclusively consisted of familicides and felony-related massacres, whereas during Wave 2 familicides dropped to roughly half of mass murders. Offenders during Wave 1 were older, more suicidal, and less likely to use firearms in their attacks (Duwe, 2004; Duwe, 2007). The change between waves is potentially a function of a post-World War II society experiencing an increase in pro-social indicators, including higher marriage rates, increased employment and compensation, and a post-Prohibition, pre-crack narcotics era. Then, during the 1960s reforms swept across the U.S. changing the social, economic, and cultural landscape. A declining economy, booming college enrollment, and rejection of the stereotypical suburban lifestyle correlate with increased divorce rates, decreased marriage rates, doubling birth rates, increased social isolation, and increased homicide and overall crime rates (Duwe, 2007). The identification of the pro-social indicators and their relationship to mass murder helps determine the social factors used in the current study for comparison to Black Swan Shootings.

Moving past Wave 2 and into the 2000s, Nekvasil and Cornell (2015) and Duwe (2007) note that mass murders have become increasingly synonymous with mass shootings. From 2005 through 2010, Nekvasil and Cornell (2015) found 22% of homicides involved two or more victims. As victim counts increase, so does the use of firearms: Nekvasil and Cornell (2015) found firearms were the primary weapon for homicides with two or more victims at least 77% of time, with firearms used in homicides with six or more victims 94% of the time. Thus, from the Duwe (2007) perspective of waves, a third wave of mass murder would seemingly be 2000-current and characterized by the dominance of firearms as the mass murder weapon of choice. Since the late 1990s, Krouse and Richardson (2015) identified mass shootings annually cluster into three types: familicide (8.5 annually), felony (8.3), and public (4.4). While more than half of mass shootings are non-public events, such as family and domestic violence, Krouse and Richardson (2015) found public mass shootings have higher casualty rates, averaging double the casualties compared to familicides and felony-related mass shootings. Among school shooting incidents, Nekvasil and Cornell (2015) also found a relatively even chance for a school shooting to have one victim (46%) versus multiple victims (54%); compared to homicides not at schools having one victim 78% of the time.

Further, Duwe (2019) more recently analyzed patterns of mass violence to understand changes in frequency over time. After controlling for population growth, Duwe (2019) found the late 1980s and 1990s to have the highest amounts of mass shootings. However, the greatest increases in events and lethality have occurred during the 2010s, which directly relates to the rise of public mass shootings. However, Duwe (2004) notes that the impact of increased public mass shootings does not impact overall mass murder trends but does significantly increase media and news attention on such events. Similarly, it is the growing public concern and increased media attention (Rocque, 2012), as well as a violent society (O'Toole, 2000), that drive fears of school shootings. These studies demonstrate the consistent increase in mass violence since the late 1990s. This research will help define a time period of analysis for the current study, and set the foundation for measuring potential temporal cycles, sprees, and contagion effects.

Offender traits and characteristics

Shifting to offender traits, perpetrators across mass violence attack types are often white males in their 30 or 40s. At least 90% of mass murderers are male (Palermo, 1997; Fox and Delateur, 2014; Meloy et al., 2001; Gill et al., 2016; Fox and Levin, 1994; Fox and Levin, 1998; Hempel and Richards, 1999), and between 34-66% of offenders are white (Palermo, 1997; Lankford, 2015; Fox and Levin, 1998; Fox and Levin, 1994; Fox and Delateur, 2014). Those percentages are on par with national demographics, as well as similar percentages of white offenders for other types of murder (Lankford, 2015). Mass shooters are similar: white males account for 63-100% of events (Bjelopera et al., 2013; Dillon, 2013; Follman et al., 2018; Lankford, 2015). According to Dillon (2013), African American, Asian, Latino, and Native American groups each comprise less than 15% of mass shooters, which is the opposite of homicide offender trends. However, as Lankford

(2015) found, the percentage of white public mass shooters (63%) is consistent with the percentage of whites in the U.S. population. School shooters and active shooters are also a majority white and male (Hobbs, 2018; Meloy et al., 2001; F.B.I., 2018; Blair and Schweit, 2014). Blair and Schweit (2014) identified six active shooter events involving females.

Mass murderers' range in age from late teens to mid 50s (Palmero, 1997), but a majority of offenders are in their 30s and 40s (Fox and Delateur, 2014; Fox and Levin, 1994; Meloy et al., 2001; Hempel and Richards, 1999). Mass shooters tend to be in their mid-30s, with an age range of 11 to 66 (Bjelopera et al., 2013; Dillon, 2013; Krouse and Richardson, 2015; Follman et al., 2018). Lankford (2015) found U.S. mass shooters on average are slightly older (34 years) than worldwide offenders (32 years). School shooters are almost exclusively teenagers, given that they typically attack their school (Rocque, 2012; Meloy et al., 2001). Active shooters are typically 14-66 years old (F.B.I., 2018; Blair and Schweit, 2014). As Fox and Delateur (2014) note, characteristics of mass murderers such as sex, race, and age are often typical for much of the general population, making their combination too generic to use as a guide or predictor of violence.

Further, mass murderers often exhibit above average rates of divorce (40% of offenders), unemployment (28%), and criminal convictions (43%); have low levels of education (24% with collegiate experience); and despite public perception, have relatively little military experience (18% of offenders) (Gill et al., 2016; Meloy et al., 2004; Hempel and Richards, 1999). Regarding employment, Hempel and Richards (1999) and Fox and Levin (1994) found impending or recent job loss is a significant agitant prior to an attack. The use and abuse of substances such as alcohol and illicit narcotics varies among mass murderers. While 62% of adolescent offenders have a history of drug and alcohol abuse, Meloy et al. (2004) found 10% of adult offenders consume alcohol immediately before their attack (Meloy et al., 2004; Meloy et al., 2001; Hempel and Richards, 1999). These are notable findings, as unemployment, poverty, and substance abuse are analytic layers in this study.

Two studies demonstrate that looking for patterns across these typical biographic traits is relatively meaningless, and that true predictive value (if any) is derived from other datasets. Lankford (2014) analyzed all global suicide attackers and found the characteristics of these offenders resemble those of conventional suicide, murder-suicide, and unconventional suicide perpetrators. Common issues include mental health problems, personal crises, coercion, fear of an approaching enemy, and hidden self-destructive urges. Lankford (2012) compared U.S. suicide terrorists to rampage, workplace, and school shooters who attempted suicide, and found that the differences between these groups, despite popular claims, were largely superficial. Most of these offenders suffered from similar problems, including social marginalization, family problems, work/school problems, and precipitating crisis events. Any perceived differences in offenders may be just that; a perception, rather than reality. As Duxbury et al. (2018) found, media portrayals of mass shooters vary by race. Duxbury et al. (2018) found White and Latino perpetrators more often have their attacks associated with mental illness than Blacks; white perpetrators are more often portrayed as sympathetic figures; and Black and Latino perpetrators are regarded as violent public threats. Further, Schildkraut et al. (2017) and

Silva and Capellan (2018) found race and victim counts are predictors for media coverage (which, in turn, has potential copycat/contagion impacts). According to Schildkraut et al. (2017), Asian perpetrators and events with higher average victim counts receive more prominent media attention. Silva and Capellan (2018) found attacks involving young, Middle Eastern, ideologically motivated offenders using multiple weapons and generating higher victim counts garner the most attention. Further, Schildkraut et al. (2017) and Silva and Capellan (2018) found attacks at locations other than schools get less coverage, with workplace events receiving the least media attention. Black Swan Shooters are expected to fall into similar patterns, featuring offenders that at a superficial level are similar to a large portion of the U.S. population.

Beyond demographics, one of the most consistent traits among mass murderers is a propensity and fascination with violence. Several studies anecdotally describe a general obsession with violence among domestic (Fox and Delateur, 2014; Meloy et al., 2001) and international (Cantor et al, 2000) offenders. Meloy et al. (2004) found 48% of adolescents and 63% of adults had obsessions of violence and war. These fantasies appear to fuel a "warrior mentality," which influences attack planning and weaponry (Meloy et al., 2004; Hempel and Richards, 1999). For school shooters, Meloy et al. (2001) describe the "classroom avenger" as a middle-class male that appears superficially normal but is considered an outcast by his peers. The offender is interested in violence but lacks a documented violent history, fantasizing about violence and meticulously planning their attack (Meloy et al., 2001).

Related to these violent tendencies are the complicated patterns of mass murderers and mental health. In some cases, mass murderers exhibit diagnosable mental illnesses prior to attack, and while these behaviors are on the mental health spectrum and are relatively consistent among this population, they are not seemingly any more prevalent for mass murderers compared to the general population (Bush and Cavanaugh, 1986). These behaviors include depression, disillusionment, paranoia, schizophrenia, long-term stress, anti-social behaviors, and isolation (Fox and Delateur, 2014; Palermo, 1997; Meloy et al., 2004; Cantor et al., 2000; Fox and Levin, 1994; Busch and Cavanaugh, 1986). Such conditions can enable offenders to misinterpret innocent actions as intentionally malicious (Fox and Delateur, 2014). However, any relationship between mass violence and mental health should be treated carefully. Skeem and Mulvey (2019) recently found that there is a relationship between serious mental illness and mass violence, but it is limited and not causal. Skeem and Mulvey (2019) note that the relationship is weaker than the public perceives, and mental illness is "neither a necessary nor sufficient condition for mass violence." Often, mass murderers are described as "loners" - individuals with tendencies to avoid interacting with others. Meloy et al. (2004) and Fox and Levin (1994) found more than 90% of adult mass murderers were described as such, despite 25% or less having a documented psychiatric history. Palermo (1997) found these external stressors create an unstable inner self, and such individuals may become more prone to such attacks.

Additionally, mass murders who are mentally ill cause more carnage. Hempel and Richards (1999) found significantly high casualty counts for psychotic offenders. Hempel et al. (2000) compared North American sudden mass assaults by a single individual to Laotian amok attacks, finding both groups of perpetrators exhibit similar degrees of social isolation and alienation, loss, depression, anger, pathological narcissism, and paranoia. Hempel et al. (2000) found that mass casualty events are not symptomatic of a specific culture, thus making them potentially not an inherently American or western phenomenon. Rather, Gill et al. (2016) describe mass murder as a unique blend of personal, political, and social factors that organize into violent action. Attackers are often characterized by a combination of unfortunate life circumstances and intense/obsessive beliefs and grievances that develop into the idea to engage in violence (Gill et al., 2016). Mental health, and descriptions of "loner" tendencies, which potentially imply social isolation, are social correlates in this study to measure across communities.

The mental health of mass murderers informs on trends for the motive of attack. It is unreasonable to merely consider all mass murderers to be irrational mad men acting impulsively and spontaneously. As Fox and Levin (1998) and Petee et al. (1997) found, most mass murderers have clear identifiable motives. Fox and Levin (1998) group mass murder and serial murder motives into five categories: power, revenge, loyalty, terror, and profit. The power motive is typically seen in the pseudo-commando offender, who turns the event scene into a war zone. Revenge attacks are typified by former employees attacking their job. Loyalty is most often seen in familicides. Profit is linked to the drug- and felony-related mass murders described in Duwe's (2007) second wave of attack trends. Terror motives are associated with groups attempting to send political messages. Phillips (2015) analyzed the lethality of terrorist attacks across 15 developed countries for

1970-2010 and found attacks conducted by organizations are far more lethal than one-perpetrator/lone wolf offenders. However, in the U.S. lone wolf attacks are often more lethal than terrorist organizations. Phillips (2015) argues this is due to robust counterterrorism strategies, making organization-based attacks more difficult to execute. However, these characteristics are not unique to the U.S., as Hilal et al. (2014) found similar patterns internationally, where the primary motivations for Chinese mass murderers were revenge and profit. In the current study, Black Swan Shootings cut across these motive typologies, lending support to the hypothesis that typologies add to the explanatory factors, but not to the predictability of such extreme attacks.

Threats are common among mass murderers but vary by type and degree. Meloy et al. (2004) and Gill et al. (2016) found a majority of offenders engaged in some form of written or verbal threats prior to their attack. Hempel and Richards (1999) found 20% of U.S. and Canadian mass murderers directly threatened their victims. The nature of threats, as Palermo (1997) indicates, can detail specific or generic frustrations and perceived wrongs with the target. These threats can be further fueled by precipitating events prior to the attack. Multiple studies indicate a specific antecedent - a major rejection, a significant loss, or perceived failure in employment, school, or personal life - plays a significant role in triggering the attack days or even hours prior (Meloy et al., 2001; Meloy et al., 2004; Fox and Levin, 1994; Petee et al., 1997). Meloy et al. (2004) found 59% of adolescent offenders and 90% of adult offenders experienced such an event prior to conducting an attack. For school shooters, threats and warnings made prior to an attack seem to cluster by school type. According to Newman and Fox (2009), middle school and high school offenders engage in attention-seeking behavior and announce threats in advance. These include clear and identifiable warning signs that other students could recognize to potentially notify authorities. Conversely, college offenders offer little to no warnings, seemingly more interested in notoriety and successfully completing the attack rather than attention (Newman and Fox, 2009). Similar to other types of mass shooters, the subsequent incident is often fueled by a rejection or disciplinary action (Meloy et al., 2001). Fox and Levin (1994) note that loss of employment as a precipitating event impacts self-esteem, finances, and social connectivity among a population (older white males) already vulnerable to increased divorce and greater isolation. For the current study, this demonstrates that communities with higher rates of socio-economic instability may have larger potential offender pools and thus greater opportunities for an attack to occur. An environment at a consistently elevated risk for factors such as unemployment or poverty may offer a greater chance to trigger a person to the violence of a Black Swan Shooting.

Aside from demographics, research indicates mass shooters do not encounter much difficulty acquiring weapons. Buchanan et al. (2015) note that despite eight of the 14 mass shootings since 2009 involving offenders with criminal histories and documented mental health problems, based on federal background checks they were still approved for firearm purchases. Fox and Delateur (2014) had similar findings: mass shooters during 2009-2013 were not prohibited by federal law from possessing firearms for mental health reasons, despite 73% of mass shooters being potentially mentally ill (Dillon, 2013).

Open source analysis from Everytown (2018) of 156 mass shootings during 2009-2016 indicates 42% of offenders exhibited warning signs prior to the attack, such as: violence, attempted violence, and threats of violence; a domestic dispute, and order of protection violations; substance abuse; and other dangerous actions (Everytown, 2018; Krouse and Richardson, 2015). However, as Fox and Delateur (2014) note, in 10 of 93 cases are such warning signs actually brought to the attention of a trained practitioner prior to an attack. Alathari et al. (2018) found relatively easy access to firearms is also a common condition for school shootings. School shootings do not typically occur in urban environments, nor as a function of school gang activity (Rocque, 2012; Twemlow et al., 2002). Rather, multiple studies have found that school shootings occur in rural and suburban areas with low crime rates (Twemlow et al., 2002; Hobbs, 2018; Meloy et al., 2001; Rocque, 2012). It is these rural areas where child access to firearms is seemingly at its greatest. Hobbs (2018) found among teenagers, access to loaded firearms without permission is most prevalent among white students in rural areas. Further, an analysis of weapons for 20 recent school shootings indicates 17 perpetrators used firearms acquired from their home (Hobbs, 2018). Looking across social correlates for the 154 U.S. school shootings during 2013-2015, Kalesan et al. (2016) found significant relationships between events and background check policies; per capita mental health expenditures; K-12 education expenditures; and urbancity. School shootings occur less often in states with background check laws, higher expenditures for mental health and K-12 education, and larger urban populations (Kalesan et al., 2016). Fridel (2019) studied all school shootings since 1998 looking for macro-level patterns across districts to determine the likelihood of an attack occurring. Fridel (2019) found high enrollments, high expenditures, greater disadvantage, and higher violent crime rates are the factors putting school districts at highest risk for a shooting. Thus, factors such as racial homogeneity, violent crime rates, and firearm availability have potential correlations to Black Swan Shootings.

It is worth mentioning the degree to which a known set of indicators are available, making such attacks potentially preventable. Lankford et al. (2019) compared 20 years of the deadliest mass shootings to the F.B.I. active shooter findings, based on violent thoughts/intent, interest in mass killing, concerning behaviors known to law enforcement, and firearm acquisition. Lankford et al. (2019) found that most of these mass shootings were preventable, in the sense that information was known about offenders prior to attack; the deadliest mass shooters exhibit more warning signs than active shooters. While this research may not account for the logistical issues of law enforcement's ability to effectively and efficiently identify and action these warning signs in advance, it confirms the potential of systematically combing through data to identify key relationships.

Attack patterns

Regarding the dynamics of a mass violence attack, the four main components to consider are the attack planning, the crime scene, casualties, and weapons. Similar to motive, mass murder should not be viewed as a singular explosion of emotion; the random public attack is the rarest of all forms (Petee et al., 1997). Rather, as several studies indicate, mass murderers plan and rehearse for days to months prior, practicing the particulars of where, when, who, and how they will conduct the attack (Fox and

Delateur, 2014; Knoll, 2012; Fox and Levin, 1994). Whether the plan is executed correctly largely depends on a combination of target availability and vulnerability, and the perpetrator's psychological and technical capabilities (Gill et al., 2016). This is true for active shooters as well; Silver et al. (2018) found 77% of active shooters spend one week or more planning the attack. Thus, these patterns of attack planning are indicative of relying on local characteristics to potentially shape the eventual attack.

For geographic patterns of mass murder and mass casualty attacks, multiple studies have demonstrated that mass shootings do not tend to occur in large, urban metropolitan areas. Fox and Levin (1998) found mass murders occur in big cities at a similar rate as single-victim homicides (34% compared to 39%, respectively); mass murders in small towns account for 43% of attacks, compared to 34% of single-victim homicides. Newman (2007) found most school shooters are from small, rural, geographically isolated communities. Newman (2007) notes that population stability for school shooter locations is extreme - either exceptionally stable (rural), or exceptionally mobile (suburbs). Further, Newman (2007) states there has never been a mass shooting at a school in a large urban community. Unlike ordinary homicide, which clusters in areas such as the deep South, mass murder distribution follows national population densities (Fox and Levin, 1998). Population density and changes over time have the potential to demonstrate unique relationships with Black Swan Shootings.

While open source research from Everytown (2018) found 63% of mass shootings occur in private homes, Bjelopera et al. (2013) found when the attack is public there is a high degree of localization and specialization; offenders target a single or closely clustered set of sites that hold significance to the offender. The top attack sites are workplaces and schools, consisting of more than half of the locations, followed by retail/shopping malls, restaurants, and their associated parking lots; public attractions including concerts, movies, and speeches; and factories and warehouses (Follman et al., 2018; Dillon, 2013; Lankford, 2015). According to Everytown (2018), 10% of attacks occur in gun-free zones. Comparatively, Lankford (2015) found attackers outside of the U.S. target military sites more often.

Active shooter attacks take place in all types of areas - small and large, urban and rural (Blair and Schweit, 2014). Active shooter attacks at commerce and education settings tend to have higher casualty counts, and a stronger connection for the shooter to the specific location (for example, a place of current or past employment). In 40% of attacks the shooter commits suicide, often at the crime scene. Further, in 10% of events the shooter initially targets family members, including estranged/former spouses and current/former girlfriends, and then moves to a public location (Blair and Schweit, 2014). Black Swan Shootings are expected to have similar patterns, where some attacks occur across mixed location types, targeting a mix of known and unknown victims. These patterns will be further analyzed in Chapter 4.

Mass murderers often culminate their attacks in suicide, much more often than serial killers - 34.7% of events versus 4.4%, respectively (Palmero, 1997; Lester, 2010). Multiple studies indicate mass murderers who commit suicide have significantly higher casualty counts than those who are captured or killed by law enforcement (Lester, 2010; Lester et al., 2004; Lankford 2015; Lankford 2015b). Krouse and Richardson (2015) and

Bjelopera et al. (2013) found more than half of attackers commit suicide, with 11-13% being killed on scene by law enforcement. Lankford (2015) found U.S. attackers die more often (61%) than worldwide public mass shooters (51%). According to Lankford (2015b) and Lankford (2015), mass murderers who commit suicide are also older, operate alone, armed with more weapons, and more likely to attack in public. However, these studies do not describe how suicide rates in the communities of these attacks effect or influence the offender; this study will measure those characteristics.

Casualty counts for mass murder vary by offender age and psychosis, with older and more psychotic offenders having significantly higher - sometimes double - casualty counts compared to younger and less psychotic mass murderers (Meloy et al., 2004). Thus, perpetrators of Black Swan Shootings in this study are typically adults and psychotic. Victims are often known to the offender and follow similar racial patterns, largely because historically most mass murders are familicides, followed by workplace and school attacks (Meloy et al., 2004; Petee et al., 1997; Fox and Levin, 1998). In cases of random victim selection, it is the location not the people that is deliberately chosen. In such cases, strangers are punished because of their perceived class membership or group association (Fox and Delateur, 2014). A mix of known and unknown victims is expected among Black Swan Shootings and will not be deliberately analyzed in this study. However, the racial homogeneity is symptomatic of community socio-economic patterns, and race will be measured here.

The preferred weapon for mass murderers is a firearm. Fox and Levin (1998) found firearms are used in more than 77% of mass murders, compared to 66% of single-victim homicides. Attackers often bring multiple weapons; according to Meloy et al. (2004), adults use 1-11 weapons (averaging 3); adolescents use 1-7 weapons (averaging 2). The prevalence of firearms is a relatively recent phenomenon, as Duwe (2004) found that prior to the 1970s firearms did not consistently correlate with mass murder. Interestingly, despite their popularity among offenders, Duwe (2004) found mass murders have lower fatality counts using firearms when compared to using bombs or fire. Of the 11 deadliest attacks occurring during 1980-2000, four involved firearms. Of the 25 deadliest attacks at the time, Duwe (2004) found firearms were used in 52% of events, compared to 67% of all mass murders. Theoretically bombs would be a better weapon choice for mass murderers; however, in practice they are unpredictable and require a higher degree of technical expertise. People are less proficient in acquiring, assembling, and using such devices (Duwe, 2004). As such, it is expected that Black Swan Shootings will most often involve adult offenders with multiple weapons. Access to such weaponry is expected to be at least partially a function of firearm availability and local laws.

Regarding weapons for mass shooters, Dillon (2013) and Follman et al. (2018) found that since the 1980s, 40-75% of attackers obtained their firearms legally. The same is true for active shooters, as Silver et al. (2018) found a majority of active shooters also obtain their firearms legally. Semi-automatic handguns are the most common firearm (Dillon, 2013), and according to Lankford (2015), U.S. attackers use multiple firearms more often than international offenders (51% vs. 22%). According to Koper (2020), high-capacity semi-automatic weapons are used in 20-58% of all firearm mass murders and cause 38-85% more fatalities; this is significantly greater than the 8%-25% usage in general firearm crimes (Dillon, 2013; Koper et al., 2017; Cannon, 2016). In fact, Koper (2020), Krouse and Richardson (2015), Dillon (2013), and Jager et al. (2018) have found relationships between these weapons and increased fatalities, injuries, and total casualties for mass murder. Of note, Cannon (2016) found that mass shooting casualties during the decade of the federal assault weapons ban (1994-2004) were 50% of those during the decade prior to the ban (119 victims vs. 241). Lankford and Silver (2019) demonstrated how increased availability to heavy artillery such as semi-automatic weapons is one of the main contributing factors to increased lethality and frequency of public mass shootings. Thus, analyzing the most severe attacks will involve a start point after the 1994 ban, and involve heavy artillery - including attacks involving multiple weapons, semi-automatic weapons, and large capacity magazines.

Personal and individual causes of mass violence

Although the focus of this study is location-centric, understanding previous research on individual causality of mass violence highlights the known trends and relative lack of individual offender predictability. The individual-centric and personality-driven causes across all types of mass violence are best understood as studies highlighting the roles of gender and mental health.

Central to the role gender plays in mass violence causality is the loss of self-identity. Multiple studies describe a male entitlement to violence among mass murderers, and attacks can be interpreted as a reaction to a threat to masculinity, integrity, and respect (Kalish and Kimmel, 2010; Ho Shon and Roberts, 2010; Kennedy-Kollar and

Charles, 2013; Tonso, 2009). Tonso (2009) describes offenders as denied or damaged in some way that violates their masculine norms, and thus feel that violence is an acceptable response to protect their identity and regain their rightful status. These denials manifest as financial, social, romantic, and psychological stressors, driven by sociocultural images, tropes, actions, and implications of masculinity that damage the offenders' perception of their masculine status in U.S. culture (Kennedy-Kollar and Charles, 2013; Tonso, 2009). The most extreme response, mass violence, is the result of the offender thinking they can change what they perceive is wrong with the world. Follman (2019) describes a strong relationship between toxic masculinity and public mass shootings. Follman (2019) found that 86% of the public mass shooters demonstrated a history of domestic violence; 32% had a history of stalking and harassment; and 50% of the attackers targeted women. Zeoli and Paruk (2019) found 31% of recent mass shooters (28 of 89) were suspected of domestic violence, and many of those were already involved in the justice system for domestic violence; 6 of the 89 had potential domestic violence-related firearm restrictions. Follman (2019) identifies clear linkages to the 'incel' subculture - a group of self-described involuntarily celibate extreme misogynists, who express violent fantasies against women. Attackers have made specific and direct references to the incel subculture in at least four public mass shootings since 2014 (Isla Vista 2014, Tallahassee 2018, Aztec 2017, and Umpqua 2015). Further, such behaviors and tendencies have also been exhibited by the perpetrators from other recent public mass shootings, including Sandy Hook, Charleston, Capital Gazette, Santa Fe, and Parkland (Mascia, 2019).

For many mass shooters, feelings of lost masculinity also relate to chronic and acute strain starting in adolescence. Levin and Madfis (2009) describe chronic strain as the long-term frustration and lack of prosocial support systems early in life leading to social isolation. Among middle school and high school males, this is driven by a failure to achieve insider status as a result of bullying, repeated rejection, humiliation, and marginalization; the persistent injustices build rage over time, as they are perceived as denials of male entitlement (Farr, 2017). Acute strain is a short-term negative event causing significant devastation. Levin and Madfis (2009) note chronic strain lay the foundation for violent behavior, and acute strain initiates the response; such a response can be the planning stage of a future attack. A mass murder is fantasized by the offender as a means to regain control (Levin and Madfis, 2009). As Farr (2017) describes, the attack is a gendered response as a culmination to their undeserved injustices, failure to achieve traditional masculinity, and overall low social standing. Thus, while these gender-based theories are individual-centric, the current study attempts to account for these indicators by measuring sex and mental health in communities experiencing Black Swan Shootings.

Shifting to individual causality of mass violence and mental health, overall there exists limited knowledge on the detection and role of mental health on mass shooting attacks. As previously mentioned, Skeem and Mulvey (2019) recently found a limited and not causal relationship between serious mental illness and mass violence. Some of this is driven by a lack of data - for offenders that do not survive, law enforcement must piece together motives and causality afterwards. Some of it, as Lankford (2016c) notes, is poor detection of mental health problems prior to the attack. In studying the challenges to detecting mental health disorders among terrorists and mass shooters, Lankford (2016c) found there is a need for better recognition and prioritization of mental health evaluation and analysis, as well as an appreciation for the role of social and situational factors. Overall, Lu and Temple (2019) have found that most mental health symptoms are unrelated to firearm violence. While hostility predicts threatening someone with a firearm, and impulsivity predicts carrying a firearm, access to firearms predicts firearm violence best, even when controlling for mental health (Lu and Temple, 2019).

When mental health does appear to relate to mass violence, there are three dimensions to examine among offenders: trauma, psychosis, and psychopathy. Langman (2009) originally identified these categories among rampage school shooters, but the archetypes also apply to perpetrators of mass violence. Traumatized attackers are typically raised in broken and abusive homes, where the parents and adult role models engage in substance abuse, criminality, and occasional physical or sexual abuse (Langman, 2009; Lankford, 2012), which leads to feelings of rejection, irritability, despondence, and violence (Palermo, 1997). Declercq and Audenaert (2011) demonstrate the lack of available attachment figures during the childhood of mass murderers fosters borderline personality disorder and furthers feelings of loneliness, depression, and predatory violence. According to Peterson and Densley (2019), most mass shooters experience early childhood trauma and exposure to violence.

Different than traumatized attackers, psychotic attackers come from intact, stable families, lacking abuse, substance abuse, or incarceration. Psychotic attackers have schizophrenia-spectrum disorders, and suffer from paranoid delusions, delusions of grandeur, clinical depression, borderline personality disturbances, rejection, recent loss, post-traumatic stress disorder, and hallucinations (Langman, 2009; Lankford and Hakim, 2011; Ferguson et al., 2011; Newman and Fox, 2009; Dutton et al., 2013; Palermo, 1997; Flannery et al., 2013). Paranoia becomes a motive for mass murder, and as alternate views are unable to be accepted, initial rejection, disrespect, and blaming of others explodes into rage and attack (Dutton et al., 2013; Palermo, 1997).

The third category of mass attackers derived from Langman (2009) is psychopathic killers. This group is neither abused nor psychotic, with no signs of abuse or dysfunction. The most commonly exhibited and identified traits of psychopathic attackers consist of narcissism, a lack of empathy and conscience, sadistic behavior, externalization of blame, frustration, extreme anger, destructive envy, a lack for identity and search for notoriety/fame/glory, and a fascination with weapons (Langman, 2009; Palmero, 1997; Lankford and Hakim, 2011; Newman and Fox, 2009; Knoll, 2012; Knoll, 2010; Fox and Levin, 1998). Unlike psychotic attackers, psychopathic attackers can often go undetected and unnoticed, despite months of planning and repeated announcement of their intentions to peers (Newman, 2007; Newman and Fox, 2009). Overall, the motivations for attack are often more practical than the psychotic attackers, including unemployment/job loss, hatred from presumed wrongs, and racial rejection (Palermo, 1997).

Overall, the research on personal and individual causes of mass violence paints a picture of an individual that lacks uniquely identifiable or distinguishable traits compared to the general public. In the aftermath of an attack it becomes clearer that planning and preparation had been occurring, as well as potential mental health risks; however, these appear to be largely unable to be discerned prior to attack (at least using current methods of detection and reporting). Thus, measuring risk at the individual level for such attacks becomes problematic. Further, Holmes and Holmes (1992) note that neurological disorders, biological dysfunctions, or other chemical imbalances are unlikely to solely cause mass violence. An appreciation of external forces should also be considered, which is where the current study seeks to bridge the gap between individual causality and community-based factors. When studying societal and community correlates to mass violence, there are spatial dynamics that influence positive social and personal conflicts. Focusing purely on mental or individual offender traits and characteristics is flawed (Rocque, 2012), and addressing violence as solely a learned behavior without considering psycho-biological predispositions is equally problematic (Palermo, 1997). As Knoll (2012) found, it is a mix of social responsibility, psychiatric assistance, cultural considerations, and media responsibility that aid research on this topic. As technology progresses and fundamentally changes social interaction, personal conflicts will continue to be difficult to identify and address (Palermo and Ross, 1999). Ultimately, it may be a mixture of external and internal mechanisms that cause these events, where a confluence of factors gels to form the precipice of an attack.

Therefore, this study advocates for a model that does not focus on the individual at all, but rather the location dynamics and elements that may precipitate or help create an environment conducive to such behavior. If risk was instead measured at the neighborhood or county (or even state) level, there exists a potential to identify more tangible and targetable patterns leading up to such an event. A location-based threat assessment approach that examines the interplay of behaviors with a social context and physical environment may provide richer insight. This study recognizes the lack of unique characteristics among offenders compared to the rest of the population and the inability to use such traits to identify future perpetrators, and instead analyzes the location/community aspects as a targetable opportunity for identifying future Black Swan Shootings.

Social factor and community causality and correlations to mass violence

There may be societal and community causes and correlations to mass violence. These can be discussed from two perspectives. The first is general trends for social factors such as media influences, social environments, and contagion effects. The second are state and community-level correlations to violence such as socioeconomic factors, firearm availability, and firearm legislation.

General social factors and trends

General trends for social factors include non-individual and societal correlates to mass violence. The non-individual factors that relate to such attacks are grouped into media portrayals and exposure to violence; geography-based pressurized and oppressive social environments; and copycat/contagion effects.

Media portrayals and exposure to violence

First, regarding the media, Duwe (2005) found at the time that while attacks were nearly as common in the 1920s as they have been since the 1960s, it is the scope and nature of national news coverage that has changed over time. High-profile attacks, which Duwe (2005) argues are the least representative of mass murder activity, receive the most attention and distort the perception of the prevalence of multiple-victim homicides. Aside from news reporting, research across academia, public surveys, and the federal government have examined the potential role violent media has on causing mass murderers and have continually found no relationship. Violent criminals on average consume less media than the average person (Salam and Stack, 2018). In Brown vs. Entertainment Merchants Association (2011), the U.S. Supreme Court noted violent video games tended to increase “hostile urges and mildly aggressive behavior” in the short-term, but found no compelling evidence these habits had any predictive value for any violent crime, let alone mass murder (Salam and Stack, 2018). A report compiled by the Congressional Research Service compared the U.S. to seven other western developed nations, finding no relationship between video games, firearm violence, and mental health (Norton, 2019). The rate of firearm violence was compared to video game revenue in Australia, Canada, France, Germany, Japan, Spain, Sweden, and the U.S. The U.S. has the highest rate of firearm violence and ownership, is second in video game revenue (Japan is first), and second in mental health disorders (Australia is first) (Norton, 2019). Pew Research Center (2015) found 49% of adults in the U.S. play video games, compared to 60% in Japan; the U.S. had 33,000 firearm fatalities in 2014, and Japan had six. In 2005, there were 30-year lows in U.S. juvenile crime despite increasing rates of youth video game playing (Salam and Stack, 2018). Any descriptions of violent media fueling firearm attacks would appear to be political rhetoric not supported with scientific evidence. Thus, media portrayals and levels of violent entertainment are not a compelling factor to analyze in this study.

More significant than playing violent video games is social isolation and exposure to violence. Tung et al. (2019) notes that qualitative studies have historically demonstrated that residents of high-crime neighborhoods will experience isolation as they have safety concerns about going outside. Quantitatively, Tung et al. (2019) have now identified that prior exposure to community violence reduces the frequency of interacting with a social network, reduces perceived social support, and increases loneliness. These are measures worth further examination in this study, as levels of violent crime and even law enforcement fatalities can serve as indicators of community safety and isolation.

Oppressive social environments

Regarding oppressive social environments and their relationship to mass violence, multiple studies identify location-based patterns of consistent instability and marginalization. Lankford and Hakim (2011) found living in an oppressive social environment is the greatest similarity between U.S. public mass shooters and Middle East suicide bombers. Conversely, there exists a relationship between high civic engagement and low mass shootings; civic engagement is typically increased through low youth populations and ethnic homogeneity (Kwon and Cabrera, 2017). An unhealthy environment can become a slippery slope that facilitates bad behavior. Among school shooters, Thompson and Kyle (2005) describes a somewhat linear path to offending, where eventual offenders are marginalized by caregivers and subsequently by peers beginning in their early years. Eventual offenders engage in social interactions that lack ethical behavior, marginalization continues, and as self-worth remains low, means of positive expression are unable to develop. A poor moral compass guides these individuals towards deviant forms of expression, to include extreme forms of violence (Thompson and Kyle, 2005).

At a broader geographic level Lankford (2016) argues the U.S. mass shooting problem may be a simple function of supply. Countries with high firearm ownership rates are more prone to public mass shootings, even when they have peaceful and healthy social indicators. Countries with high ordinary homicide risk, including Venezuela, Nigeria, and Mexico, have low rates of mass shootings (Lankford, 2016). According to Lankford (2016), the U.S. has 5% of the world's population, and 31% of its public mass shooters. Interestingly, Altheimer and Boswell (2012) also found that economic inequality had a stronger relationship with homicide rates than firearm availability; suggesting that socioeconomic factors may have a relationship to firearm violence. These studies highlight the impact social environments have on mass shooting activity. These are valuable patterns for identifying common locations and areas for Black Swan Shootings.

Copycat and contagion effects

Research on copycat and contagion effects of mass murder has been historically mixed, with recent studies indicating such phenomena do exist. While Fox and Delateur (2014) find no evidence of copycat or contagion effects, suggesting such claims are merely a collection of anecdotes, and mass murder occurs independent of behavioral modeling, other studies are less certain. Duwe (2000) argues high-profile cases shape perceptions, but since there are relatively few events it is difficult to measure any short- or long-term effects on such a small set of cases. Duwe (2007) found that if contagion exists, it is a rare phenomenon. Of the 909 events Duwe (2007) studied, five have demonstrable influence from a previous mass murder. What is more probable is that of the 4-5 cases annually that receive national, mainstream media coverage, and any temporal clustering frames the distinctly unique events as an epidemic (Duwe, 2007). The contagion may be media coverage of the events, rather than the attacks themselves. Also, there may be national forces rather than community-specific effects that are driving the occurrence of these events.

However, since the publication of Duwe (2007), public mass shootings are occurring at a historically deadly and fast pace. As Lankford and Silver (2019) note, the increase in occurrence and lethality are also correlated with increases in fame-seeking offenders. Younger generations are valuing fame over traditional ideals such as family, marriage, and professional careers (Lankford and Silver, 2019). Lankford and Silver (2019) note the blurred lines between fame and infamy, and the premium being placed on attention. Similarly, Cantor et al. (1999) highlights the necessity of media coverage for a chain of copycat causation to occur. Cantor et al. (1999) identified offender modeling in four of seven international mass murders, where the news reporting resonated with future offenders for up to ten years, as seen by the linkages made in perpetrator statements. Lankford and Madfis (2018) also note the role media coverage plays in copycat and contagion effects, with increased media attention creating a celebrity status for a small population drawn to such deviance. As Lankford and Madfis (2018) note: “psychologically healthy people do not commit mass shootings based on what they read or see in the news, but there are troubled and at-risk individuals who respond very differently...Even dangerous or distasteful products, people, and experiences almost always garner more interest if they are widely advertised than not advertised at all." Further, as Langman (2018) notes: "for people already at risk or on a path toward violence, however, external influences in the form of other mass attacks may be a factor in spurring them on toward committing their own attack" (p. 1).

Internationally, Carcach et al. (2002) found no long-term effects of a single mass murder attack on general homicide rates; if anything, the Port Arthur massacre in Australia ushered in long-term homicide and firearm violence declines. But in the immediate aftermath, Carcach et al. (2002) did identify a five-day spike in homicides, directly attributed to the attack.

Similarly, Towers at el (2015) examined U.S. mass killings and identified a statistically significant 13-day increased probability for follow-on copycat mass attacks. It is a temporary increase that causes 0.30 new events (Towers et al., 2015). Jetter and Walker (2018) examined the amount of mass shooting coverage on ABC World News Tonight for 2006-2016, finding a positive statistically significant relationship between the amount of coverage and future attacks. According to Jetter and Walker (2018), this effect lasts for 4-10 days after an initial mass shooting, and this window explains more than half of all mass shootings. Related, Kissner (2016) found active shooter incidents are temporarily contagious, where new attacks are a function of activity that occurred within the preceding two-weeks. Thus, recent literature suggests that if contagion effects exist, they result in increased probabilities for new mass shootings in the 4-14 days after an attack.

Similarly, there are elements of spatial contagion to examine. Multiple studies have described crime movement as the spread of disease or infection. Zeoli et al. (2014) studied spatial and temporal movement of homicides in Newark, New Jersey, where homicides are the infectious disease and firearms and gangs are the infectious agents. Zeoli et al. (2014) describe a single strand/cluster of activity in the city center that slowly spread over 20 years to the most contagious areas, where other parts of the city remained immune - seemingly due to their lack of location similarity to the origin area. Messner et al. (1999), studying homicides in St. Louis found similar results of spatial spread, noting that such diffusion occurred nonrandomly, and affluent, rural, and agricultural areas served as barriers to spread. Johnson et al. (2007) applied epidemiological techniques for disease contagion to residential burglaries and found in the aftermath of a burglary, there is an elevated risk within 200 meters and 14 days for future events to occur, noting this is partially a function of target density. This study uses similar epidemiology target clustering models to identify spatial contagion effects for Black Swan Shootings. This has applications for understanding spread and decay across counties rather than neighborhoods, to include a potential smoothing effect for attack probabilities over space and time.

Further, related to mass violence and social environments, D'Anna (2016) examined the clusters of supporters for firearm restrictions and supporters of firearm ownership, compared to the locations of mass shootings. Using social media self-reporting, D'Anna (2016) found supporters of the firearm rights are densely clustered in non-mass shooting areas of the country, while firearm control supporters are more dispersed across mass shooting hotspots. Firearms rights supporters live significantly farther away from mass shootings than firearm control supporters. This study is not without its flaws, as it does not consider any temporal factors for a mass shooting, as well as relying on social media location self-reporting. However, it connects disparate concepts in this literature review of mass shootings, demographic clustering, and attitudes about firearms. Thus, the current study builds upon this concept by adding more social factors to test over time.

Related to copycat offending is fame-seeking attacks. Silva and Greene-Colozzi (2019) found that since the early 2000s there has been an increase in fame-seeking mass shooters. The perpetrators are often young white males, targeting schools, and exhibiting signs of mental illness, suicide tendencies, and grandiose behaviors. According to Silva and Greene-Colozzi (2019) and Langman (2018), such attacks receive significant media attention, which negatively reinforces the behaviors and motivations, and promotes them to a celebrity status. Langman (2018) highlights that many mass shooters have role models leading up to their attacks. These role models "inspire imitation" as an exalted figure (Langman, 2018). Similar to Duwe (2007), Larkin (2009) examines the effects of a broader influence, finding the influence of the Columbine High School attack on April 20, 1999 was a start point for a generation of attacks. Larkin (2009) found in every school shooting analyzed there were admitted links or direct evidence of Columbine influences. In this model, Columbine is the start point in four ways. First, it is the paradigm event for planning and executing a mass school shooting. Second, it inspired subsequent attackers to avenge wrongs and humiliations through violence. Third, it created the baseline for casualty counts, against which all future attacks would be measured. Fourth, its perpetrators were vaulted to mythical status among deviant subcultures (Larkin, 2009). As such, this fits the Duwe (2000) perspective on high-profile attacks shaping perceptions - driven by the media, and consumed by the general public, policymakers, and eventual offenders. Thus, as it pertains to the current study, if copycat behavior exists, it is evidenced in long term attack similarities, patterning, and offender modeling.

Community-level causes and correlations to mass violence

Different from societal and community causal relationships to mass violence, there are also macro-geographic correlations to mass violence. These include socioeconomic factors, firearm availability, and firearm legislation.

Socioeconomic factors

Previous research relevant to the current study also examines location correlates with socio-economic factors. This research is relevant because it demonstrates the importance of neighborhood-level clustering of non-crime factors. First is the notion that non-crime factors cluster. Multiple studies find that similar to the Shaw and McKay (1942) theory of social disorganization, neighborhood traits cluster - and thus further structural disadvantage - by physical status, economic status, racial composition, residential mobility, and level of social control, which contribute to levels of violent crime (Bernard et al., 2015; Peterson and Krivo, 2005; Sampson and Wilson, 1995).

In addition to race, and in the context of the current study, McPhedran and Baker (2011) found social change is a cause of stress. Individual stressors that impact socio-economic status, in particular divorce and unemployment, may help understand mass shooting causality. Mass shooting activity potentially temporally clusters around periods of poor economic health and high unemployment and are nonexistent during healthy economic times (McPhedran and Baker, 2011). As such, if mass shootings (and general firearm violence) are seen as a public health problem, the U.S. largely stands alone. Wintemute (2015) found U.S. mortality rates remain relatively unchanged; despite homicide rates decreasing, suicide is increasing - and the overwhelming majority of both metrics are driven by the use of firearms. Risk of firearm victimization is not distributed evenly across the population; Wintemute (2015) notes that the burden of illness clusters in demographic subsets. Thus, measuring public health indicators (economic and social) across space and time serve as potential predictors of Black Swan Shootings.

Further, studies on socio-economic correlates to crime are largely based on youth, race, and poverty. Early studies, such as Blau and Blau (1982), demonstrated high levels of violence related to proportions of African American populations, high poverty, and disproportionate divorce rates. And more recently, studies such as Robinson et al. (2009) demonstrate high rates of high school dropouts, unemployment, and racial clustering can account for as much 90% of the variation in homicide rates across communities. Regarding youth, Florida (2011) found a positive correlation at the state level for firearm fatalities and high school students carrying weapons at school (0.54). Burdick-Will (2018) found exposure to neighborhood violence affects youth school behavior and engagement; when peer involvement in violence is high, it distorts a student’s trust, discipline, and perceived safety in school. Related, Gurney and Teproff (2016) found in Miami, Florida approximately 20 zip codes account for a majority of youth violence crimes. Among potential victims, a local coalition identified 2,000 children at the highest risk for firearm violence, based on low attendance, classroom behavioral issues, and poor test scores for those living in the hotspots. Similarly, Papachristos et al. (2015) found that risky social networks relate to firearm violence clustering; as exposure to violence increases, risk of victimization dramatically increases as well. In Chicago, 70% of non-fatal firearm victims come from co-offending networks of 6% of the population. Papachristos et al. (2015) notes that these networks are most prevalent for minority youth males. While these studies seek to explain the relationships between socioeconomics and general firearm violence, they also establish the foundation for conditions that may be conducive for violence to occur in a community.

Socio-economic correlates for mental illness and public opinion are less researched but relevant to the current study. While Florida (2011) did not find relatively strong correlations for community mental illness rates when compared to economic, legislative, and social factors, other studies have found more nuanced relationships. Elbogen and Johnson (2009) found that mental illness alone does not predict violence (which is consistent with studies from earlier sections on the prevalence, trends, and causes of mass violence), but mental illness combined with a violent history, substance abuse, demographic traits, and economic factors demonstrates a stronger pattern of violence. This further highlights the notion of converging factors creating a greater opportunity for mass violence to occur.

Markowiak et al. (2018) identified social health correlates as risk factors for mass shooting attacks. Using violent crime data, state firearm laws, socioeconomic factors, and community behavioral measures, Markowiak et al. (2018) found communities at higher risk for mass shootings had with the strictest state firearm laws, were less social, had less leisure time activity, were less rural, younger, had high rates of overcrowding/lack of utilities, and had high income inequalities. Communities without mass shootings had higher mental health professionals per capita (Markowiak et al., 2018). Markowiak et al. (2018) concluded communities should focus on safeguarding citizens with access to mental health professionals, promoting socialization, increasing public space usage, and addressing socioeconomic inequality. Related, two studies from Kwon and Cabrera (2017, 2019) highlight the relationship between mass shootings and socioeconomics. First, Kwon and Cabrera (2017) found a relationship between economic instability and mass shootings. This relationship appears to be driven by high poverty and single-parent households. Second, Kwon and Cabrera (2019) found counties with growing levels of economic inequality are more likely to experience mass shootings. Thus, these studies directly apply to the social factors analyzed in the current study. All three studies pave the way for demonstrating the location correlates for violence, mental health, and socioeconomics that can play a role in the risk/occurrence of such activity.

Firearm availability

Prior research on firearm availability demonstrates the relationship of access to weapons and firearm cultural conditions to violence – and in particular, mass violence. Multiple studies on firearm availability rely on public surveys. Such research indicates declines in the percentage of households and adults with firearms but increases in number of firearms per owner (Cook and Ludwig, 1997; Krouse, 2012; Hepburn et al., 2007). Krouse (2012) estimated civilian ownership has doubled since 1968, and Hepburn et al. (2007) determined the U.S. population has at least one firearm per adult. Ownership appears to cluster by individuals and communities, indicative of a firearm culture existing in specific pockets throughout the country. Hepburn et al. (2007) found roughly half of firearm owners possess four or more weapons. Kalesan et al. (2015) notes firearm ownership is 2.25 greater within firearm cultures. Thus, surveys have demonstrated that firearm ownership across the country is high, but even higher in specific regions and communities.

Different from survey data, other studies sought to develop and validate more accurate proxy measurements. In particular, Kleck (2004) found that using the percentage of suicides committed with firearms is the single best indicator of firearm ownership. This measure had a 0.87 correlation with firearm surveys across large cities, 0.92 across states, and 0.95 across countries (Kleck, 2004). This is further supported by Hemenway and Miller (2000), who used percentage of suicides committed with a firearm to determine a strong, statistically significant relationship between firearm availability and homicide rates. This proxy is used in the current study because suicide data is more readily available than public surveys.

There is a clear, strong relationship between firearm availability and fatalities, which include homicide and suicide. Lester (1993), Lemieux (2014), and Anglemyer et al. (2014) are among those who found that higher rates of firearm availability lead to higher firearm death rates. This finding is almost universally consistent among studies examining homicide and suicide across gender (Hepburn and Hemenway, 2004; Miller et al., 2007; Siegel and Rothman, 2016; Rodríguez and Hempstead, 2011); age (Miller et al., 2002; Miller et al., 2007); mental health factors (Hemenway and Miller, 2000);

socioeconomic differences (Miller et al., 2006); income levels (Hepburn and Hemenway, 2004); legislation (Duggan, 2001); at city, state, region, and country boundaries (Shenassa et al., 2006; Hepburn and Hemenway, 2004; Miller et al., 2002b; Briggs and Tabarrok, 2013; Siegel and Rothman, 2016; Miller et al., 2007; Azrael et al., 2004; Miller et al., 2002); internationally (Lewiecki and Miller, 2013; Killias et al., 2001); and weapon type (Zimring, 1968). Firearm availability is often the single best predictor of firearm death rates (Lemieux, 2014).

The patterns hold up internationally, supporting the notion that the U.S. is a microcosm, not an exception, of violent human behavior. Internationally, Killias (1993) examined 11 European countries, the U.S., and Canada, Lemieux (2014) studied the U.S. and 25 other western democracies, and Altheimer and Boswell (2012) studied 43 countries for firearm availability and homicide rates. In each case, there exists positive correlations between firearm ownership and national homicide and suicide rates. Thus, it is reasonable to conclude that the presence of firearms in a community would increase the opportunity for a Black Swan Shooting to occur.

Anisin (2019) found that a culture of fear spawned by Columbine (1999) and the 9/11 terrorist attacks (2001) has Americans arming themselves at a historically high rate. This has not drastically changed the percentage of the population that owns firearms; rather, it has increased the number of weapons an existing firearm owner has (Kalesan et al., 2015). Fear, and in particular fear of mass shootings, has led to this armament (Anisin, 2019). Not all mass shootings spur the public to arms, but it is the combination of fear and viral media attention from specific events that appears to be fueling this increase. And among those events, it is a mix of the offender's characteristics, the attack location type, and the fatality count that spark fear (Anisin, 2019). Thus, in the current study firearm ownership is a social factor compared to Black Swan Shootings.

Related to firearm ownership is firearm purchasing. Studdert et al. (2017) analyzed firearm purchases in the aftermath of two high-profile mass public shootings: Newtown, Connecticut in 2012, and San Bernardino, California in 2015. In the six weeks after each attack, there were increases in firearm acquisitions of 53% and 41%, respectively. The largest increases were seen among Caucasian and Hispanic populations, people who did not previously own a firearm, and residents in adjacent cities and neighborhoods of the attacks. In particular, after San Bernadino, firearm acquisition jumped 85% in adjacent areas, compared to 35% across California (Studdert et al., 2017). It is worth noting that these sudden, brief increases in firearm acquisition may correlate with the aforementioned public mass shooting contagion effects. Nationally, the FBI has reported an increasing amount of firearm background checks; 2019 was the highest year for such checks since the FBI began tracking in 1998 (F.B.I., 2019). While mass shootings have also increased during this time period, it is unclear how strong the correlation is between both measures.

Multiple recent studies demonstrate how firearm availability is a factor in creating an environment conducive for a mass shooting. Reeping et al. (2019) found a 10% increase in state firearm ownership correlates with a 35% higher mass shooting rate. Towers et al. (2015) found that states with high firearm ownership are significantly associated with state-level mass shootings and school shootings. Firearm ownership is found to be more significant and outweighing any relationships for mass shootings with mental health factors or firearm legislative strength. Lankford and Silver (2019) note that the ability to obtain high-powered weaponry has increased over time. Specific to semi-automatic rifles and assault weapons, the number of these manufactured annually in the U.S. has grown from 1 million in 1986 to 4 million in 2016. Despite temporary limits to these weapons from the Federal Assault Weapons Ban during 1994-2004, sales and production quickly bounced back and increased. Further, with supply spiking, prices have dropped, significantly increasing opportunities for acquisition (Lankford and Silver, 2019).

Firearm legislation

Research on the ability of U.S. firearm legislation to impact crime has become more convincing over time. Early studies often found mixed or no relationship between firearm laws and total crime. However, as research methods have improved at isolating variables and controlling for noise, significant relationships have been identified. The number of and types of firearm laws are proxies for attitudes on and accessibility of weapons in different communities. As such, previous research related to firearm legislation applicable to the current study is divided into two categories: the impacts of firearm legislation on crime; and the impacts of firearm legislation on mass violence.

Before discussing the research, it is worth defining some of the terms commonly associated with firearm legislation:

1. Duty to Retreat (DTR): making an attempt to flee from an event prior to using deadly force.

2. Permit to Purchase (PTP): requiring purchasers to be licensed and have a law enforcement background check conducted prior to firearm ownership.
3. Right to Carry (RTC) and Shall Issue (SI): mandating/permitting weapon possession via permit is granted to all qualified applicants unless restricted by another statute.
4. Stand Your Ground (SYG): prosecutorial immunity from self-defense with deadly force. (Isenstein, 2015)

Studies on the relationship between domestic firearm legislation and crime cut across general crime, violent crime, and fatal crime. Studies rarely find direct evidence of firearm-related legislation impacting total crime rates. When Moorhouse and Wanner (2006) examined 30 components of state-level firearm laws, the significant relationship was not with crime and laws, it was that high crime rates generated more political support for stricter firearm laws. Dezhbakhsh and Rubin (1998) found that laws increasing conceal carry in a state result in small reductions in murder, increases in robbery, and county-to-county variation among other crime types. Ayres and Donohue (2003) found adoption of SI firearm laws have the greatest effects on decreasing crime in four distinct ways: 26 states should repeal SI laws; 19 states should never adopt; six states should continue current laws; and California should adopt a SI law. The takeaway: at the general crime level, firearm legislation may have impacts, but it is inconsistent and not convincing.

Research on the effects of firearm laws on violent crime is clearer. Vittes et al. (2012) found laws on background checks for firearm ownership made it more difficult for violent offenders to obtain firearms. Webster and Wintemute (2015) found that legislation designed to prevent high-risk populations from owning firearms, including PTP, background checks, dealer oversight, and lost/stolen reporting requirements, are negatively associated with criminal access to firearms. Jung and Jason (1988) found laws for stricter firearm possession in public have temporary impacts on reducing robberies and assaults committed with a firearm. Simonetti et al. (2015) found evidence of stricter state firearm legislation associated with lower hospital discharge rates for nonfatal firearm injuries. Donohue et al. (2018) found statistically significant higher violent crime rates in states with RTC laws. Over time, Donohue et al. (2018) note violent crime is expected to be 13-15% higher in such states, which has the potential to double the prison population in those states. Aneja et al. (2014) identified a 33% increase in firearm assaults, as well as increases in sexual assault, robbery, and murder rates, for states with RTC laws. Regarding domestic violence, Webster and Wintemute (2015) found strict domestic violence restraining orders are associated with lower rates of violence. Vigdor and Mercy (2006) found 7% reductions in homicide rates after states passed legislation on domestic violence firearm confiscation. This research indicates that specific laws can have targeted effects on violent crime - including RTC laws increase crime, background checks decrease crime - which informs the current study's location correlations.

The impact of firearm legislation on fatalities (homicide and suicide) is most telling. Multiple studies at the state level demonstrate that higher numbers of firearm restriction laws are associated with lower rates of suicide, homicide, and total fatalities (Fleegler et al., 2013; Geisel et al., 1969; Florida, 2011; Lester and Murrell, 1982;

Anestis and Anestis, 2015). Kaufman et al. (2018) had similar findings, adding that even counties of states with weak firearm laws had lower homicide rates when the adjacent states had strong laws. In some cases, such as Parsons and Weigend (2016), the relationship is quite strong; the ten states with the weakest firearms laws have 3.2 times greater levels of homicide than the ten states with the strongest laws. Isenstein (2015) found five of the six lowest firearm fatality rates clustered in the Northeast, and four of the six highest firearm fatality rates clustered in the southeast. In other cases, such as Kwon et al. (1997) and Boor and Bair (1990), the same laws-to-fatalities connection is found. However, it is worth noting that the relationships are relatively mild when compared to the relationship between firearm fatalities and socioeconomic measures such as poverty, unemployment, divorce rates, and alcohol consumption. The findings of such correlates from Kwon et al. (1997) and Boor and Bair (1990) determine which social factors to consider for spatial correlates to Black Swan Shootings.

Examining specific types of firearm legislation to firearm fatalities, findings on the effects of comprehensive background checks for firearm purchases and reductions in firearm fatalities are supported by multiple studies (Ruddell and May, 2005; Bisakha and Panjamapirom, 2012; Lee et al., 2016; Ludwig and Cook, 2000). Related to background checks are waiting periods and permit to purchase processes for firearm ownership. Research on the effects of such legislation on firearm fatalities has consistently demonstrated a negative relationship (Anestis and Anestis, 2015; Crifasi et al., 2018; Lee et al., 2016). Luca et al. (2017) found that laws using a waiting period of several days resulted in a 17% reduction in firearm homicides and extrapolating nationwide such a waiting period could prevent approximately 750 firearm homicides annually - without coupling with any other legislation. Crifasi et al. (2018) found PTP laws are associated with a 14% reduction in firearm homicides across large, urban counties. Webster et al. (2014) and Crifasi et al. (2015) found 16-23% increases in firearm suicides and homicides in Missouri, after repealing a PTP law; this equated to an additional 55-63 homicides annually. Connecticut experienced a 15% reduction in firearm suicides when implementing such a law (Crifasi et al., 2015), and Rudolph et al. (2015) found the PTP law was associated with a 40% reduction in Connecticut firearm homicides in the ten years after its passage. Swanson et al. (2016) examined the impact of Connecticut's law for extreme risk protection orders, which removed firearms from high-risk individuals, finding that more suicides would have occurred without the law, estimating that 10-20 firearm seizures occurred for one suicide averted. Webster et al. (2002) found bans of specific "Saturday Night Special" handguns in Maryland resulted in 6.8-11.5% reductions in firearm homicide rates.

Internationally, the patterns are similar. Santaella-Tenorio et al. (2016) examined 130 studies from ten countries and found legislation restricting firearm access is linked to reductions in firearm fatalities, with the laws associated with the most significant reductions being background checks and safe storage. Cantor and Slater (1995) found a 28-day waiting period prior to firearm purchase in Australia resulted in a statistically significant drop in firearm suicide rates. In Colombia, Vecino-Ortiz and Guzman-Tordecilla (2019) found cities that enacted permanent firearm bans saw 22% reductions in monthly firearm-related mortality rates. The takeaway here: stricter laws related to licensing and permit to purchase; domestic violence and high-risk offender restrictions; waiting periods; and comprehensive and universal background checks reduce firearm fatalities.

Conversely, legislation related to SI requirements, RTC, concealed carry licensing, and SYG legislation is often associated with increased or higher fatality rates. Multiple studies have found at a minimum, SI laws do not reduce firearm homicides and more often are actually associated with an increase (McDowall et al., 1995; Ludwig, 1998; Rosengart et al., 2005). Siegel et al. (2017) found increased rates for homicides (6.5%), firearm homicides (8.6%), and handgun homicides (10.6%) in states with SI laws. Similarly, studies such as Crifasi et al. (2018) demonstrate that RTC and SYG laws are associated with increased firearm fatality rates. Isenstein (2015) compared states with SYG versus DTR laws and found DTR states had a 1.52 lower homicide rate. In Florida, Humphreys et al. (2016) found the introduction of a SYG law was related to a sudden 24.4% increase to the monthly homicide rate, and a 31.6% increase in firearm homicides; notably, homicide rates were on decreasing trend prior to the law's implementation. Further, Rand Corporation (2018) found support that SYG laws may increase violent crime and found no qualifying studies that SYG laws have the ability to decrease violent crime.

Thus, laws covering SI, RTC, and SYG have not worked to reduce fatalities; in fact, they seem to increase homicides. However, as Seitz (1972) noted, legislation-related results historically affect Caucasian culture the most. Also, more restrictive places overall have lower homicide rates. Hence, there is great value for the current study to examine firearm legislation and its relationship to Black Swan Shootings, but also consider the interplay of these laws with socio-economic location correlates.

Comparatively, research on firearm legislation and the impacts on mass shootings is not as robust. However, the volume of academic studies is increasing, as such attacks are more prevalent (and federally funding is becoming available). Most studies that measure associations of RTC, gun-free zones, and child protection laws to mass shootings have found no impact. RTC laws are theoretically designed to increase the armed populace, and serve as a deterrent for public mass shooters, as there becomes a larger (albeit unknown) group of people available to intercede in an attack. In reality, Duwe et al. (2002) and Webster et al. (2016) have found the opposite - there is no evidence RTC laws have any impact whatsoever on public mass shootings. Further, theoretically "gun-free" zones (such as schools) would be attractive target sites for public mass shooters as they lack guardianship. However, Webster et al. (2016) found 12-13% of mass shootings occur in such areas. Allowing or increasing firearm possession in such protected places is not expected to deter any activity, and realistically expected to increase and escalate violence (Webster et al., 2016). Kleck (2009) analyzed the impact of legislation introduced after the 1999 Columbine attack. Kleck (2009) examined gun show restrictions, child access prevention, and assault weapons bans, and at the time found no impact of such legislation on mass shootings. However, Kleck (2009) did find support that those measures significantly help prevent ordinary firearm violence.

Since 2018, some of the traditional lag in U.S. policy changes in the immediate aftermath of a mass shooting has changed. After some of the deadliest attacks in history took place in Parkland, Florida and Las Vegas, Nevada, lawmakers across the country began passing firearm laws at an unprecedented rate. There were 50 new laws were passed to restrict firearm access, ban bump stocks, and establish extreme risk protection orders (Vasilogambros, 2018). Unfortunately, when the U.S. has attempted to change laws after a mass shooting, historically the new legislation misses the mark. Keneally (2019) found that when U.S. states pass reactive legislation in the aftermath of a significant mass shooting, the laws implemented would not have even prevented the attack they are reacting to. Keneally (2019) notes that often the focus is to minimize illegal firearm purchases; however, all attacks involving ten or more fatalities over the past 20 years involved legally purchased weapons (and often by the shooters directly). Instead, legislative bans on bump stocks, or implementing extreme risk protection orders, would have the potential to have prevented previous attacks. A bump stock was used in Las Vegas to alter semi-automatic weapons to fire at a speed similar to full-automatic ones. Extreme risk protection orders allow for law enforcement, family, and the community to report concerns for a firearm owner, and a judge to temporarily remove firearms from them. Such a policy had the potential to disrupt the attacks in Columbine, Virginia Tech, Sandy Hook, and Aurora (Keneally, 2019). In fact, as Laqueur and Wintemute (2019) and Wintemute et al. (2019) found, after California implemented such red flag laws (known as Gun Violence Restraining Orders), there have been 21 cases of prevented mass shootings. Among 159 records received and analyzed, these 21 cases were mostly males making explicit threats and owning firearms, and the law was able to authorize the removal of weapons, where no such mass shooting has occurred involving those individuals (Wintemute et al., 2019).

There are multiple recent studies that have found legislative impact on mass shootings. Analyzing state firearm laws and the rate of mass shootings, Reeping et al. (2019) examined the restrictiveness/permissiveness of each state on a scale of 0-100 from 1998-2015, and found a relationship between permissive gun laws, higher firearm ownership, and increased rates of mass shootings. Webster et al. (2020) examined the relationships between mass shooting fatalities and multiple firearm laws. Webster et al. (2020) found handgun purchaser licensing laws and large-capacity magazine bans are associated with significant reductions in mass shootings. They also found comprehensive background checks, assault weapons bans, and de-regulating civilian concealed carrying were not related to mass shooting activity. Further, despite finding 28% of mass shootings had a connection to domestic violence, Webster et al. (2020) found no support that laws restricting access to firearms for domestic violence offenders would have prevented those attacks. Of note, Webster et al. (2020) describe how licensing laws reduce firearm availability to those at risk and have positive effects on mass shootings and general firearm violence. Guis (2015) measured associations to state and federal assault weapons bans to public mass shootings, finding such bans had negative effects on fatalities. Federal bans had negative effects on public mass shooting injuries. Addressing the impact of banning high capacity semi-automatic firearms on mass shootings, Koper (2020) analyzed multiple contemporary studies and found strong support. Koper (2020) found 20-58% of mass murders involve semi-automatic weapons (which include assault rifles); such attacks result in more fatalities and injuries than such attacks with other firearms. Further, after the federal assault weapons ban expired in 2004, mass shootings involving such weapons increased. Koper (2020) suggests state-level restrictions on such weapons and large capacity magazines has the potential to reduce mass shootings by potentially 11-15%.

An analysis in the Los Angeles Times using attacks from The Violence Project demonstrates how legislation has the potential to prevent mass shootings. Mukherjee (2020) analyzed the 167 mass shooting attacks captured in The Violence Project database since 1966 and compared then to give groups of firearm regulations: bans on straw purchases, safe storage requirements, assault weapons bans, mandatory background checks, and red flag laws. Overall, Mukherjee (2020) found that had all of these measures been in place federally, 146 of the 167 mass shootings could have potentially been prevented; this includes preventing all but one attack in the past five years. Further, the potential prevention of mass shootings varies across the law types. Mukherjee (2020) found these potential preventions skew towards red flag laws, which had the potential to prevent 141 mass shootings. Next highest are assault weapons bans, with the potential to prevent 38 attacks. The other three law types, mandatory background checks, safe storage requirements, and straw purchase bans (which is buying a firearm for someone else), would have potentially prevented 16, 14, and 5 attacks, respectively (Mukherjee, 2020).

Finally, Lemieux (2014) examined a mix of cultural and legislative impacts on mass shootings. By comparing international and domestic mass shooting fatalities to cultural artifacts surrounding firearms and firearm laws, Lemieux (2014) found both concepts have significant impacts on mass shootings. The cultural perspective is associated with increased mass shooting fatalities, and strict legislation is associated with decreased mass shooting fatalities (Lemieux, 2014). As such, it is this more nuanced combination of social and legal factors that each seemingly weigh into the location-based correlations of mass shootings, lending credence to the idea of a mixed model for measuring them in comparison to Black Swan Shootings.

Internationally, among western nations most similar to the U.S., it has become somewhat standard practice for firearm legislation to significantly tighten in the immediate aftermath of a catastrophic public mass shooting. Unlike these countries, the U.S. does not make sudden drastic changes to its laws, despite the frequency and intensity of these events occurring in the U.S. (Covucci, 2016), as well as the growing public sentiment for crime reduction policies (Cohn et al., 2013). In Canada, after a mass shooting in 1989 in Montreal killed 14 people, new legislation included a twenty-eight-day waiting period for purchases; mandatory safety training courses; more detailed background checks; bans on large-capacity magazines; and bans or greater restrictions on military-style firearms and ammunition (Masters, 2017). In Australia, within two weeks after the mass shooting at Port Arthur in 1996 that killed 35 people, firearm regulations including bans and buybacks of different weapon types, as well as licensing, registration, storage, and training requirements were implemented (McPhedran and Baker, 2011; Reuter and Mouzos, 2003). After another high-profile shooting in Melbourne in 2002, firearm laws in Australia were further tightened (Masters, 2017). In the United Kingdom, a mass shooting spree at Hungerford in 1987 killing 16 people triggered a law banning semi-automatic rifles and increased registration requirements for other firearms. Then, after a mass shooting at Dunblane, Scotland in 1996 killed 17 people, the country banned handguns and implemented a firearm buyback program (Masters, 2017). In Norway, historically firearm ownership is among the highest in the world, while firearm homicide rates are among the world's lowest. However, after the 2011 shooting in Oslo that killed 77 people, the Norwegian government implemented age, use, and license requirements, as well as considering (but ultimately not implementing) bans on pistols and semi-automatic weapons (Masters, 2017). The global trend of reactionary legislative changes seems to correlate with a low numbers of future mass public shootings.

However, research regarding the effects of such legislation in foreign countries is relatively sparse. Sherman (2000) notes the firearm buyback program in the U.K. has not been formally evaluated, and despite research on the Australian buyback, the results are not generalizable to the U.S.; the relative firearm density and firearm homicide rates have always been distinctly higher for the U.S. Most academic research on legislative effects on foreign mass shootings ultimately examines the 1996 Port Arthur shooting. The findings are decidedly mixed. Among studies that have demonstrated no impact, research shows firearm homicides, suicides, and accidental fatalities have not significantly changed despite increased regulations on possession, registration, storage, and waiting periods (Baker and McPhedran, 2007; Carcach et al., 2002; McPhedran and Baker, 2011). And while some studies have credited legislative and public policy changes with a lack of further mass shootings (Carcach et al., 2002), other studies at the time found no difference for mass shootings between Australia and neighboring countries (McPhedran and Baker, 2011). Post Port Arthur attack, Reuter and Mouzos (2003) suggest modest effects, where suicide rates did not change but robbery rates increased (and in both cases weapon type changed to less firearms), and firearm homicides continued their slight decline prior to the attack. Peters (2013) and the Library of Congress (2015) found the firearm fatality rate in Australia was cut by more than 50% of what it was prior to Port Arthur, and a fraction of the U.S. rate. Chapman et al. (2006) and Chapman et al. (2016) noted there were 13 mass shootings in Australia in the 18 years prior to Port Arthur, and none in the 20 years after. Chapman et al. (2006) and Chapman et al. (2016) argue the aforementioned declines in firearm fatalities were accelerated by the post-Port Arthur legislation, citing the weapons bans as potentially effective means to reduce mass shootings, firearm homicides, and firearm suicides.

Thus, in this study understanding the effects of firearm legislation on the locations of Black Swan Shootings can be based on the number of laws in a community. It appears the higher the number of firearm access prevention and safety laws that exist, the lower the probability for violence. Further, this literature review indicates that communities with the most restrictive laws have the lowest propensity for violence, but that does not always correlate with the type of rare mass shooting attacks analyzed here.

Summary

In conclusion, this literature review highlights mass violence trends, person-centric causes of mass violence, and social factor correlations to mass violence. The research on mass violence trends indicates a clear, significant increase in such attacks during the 2000s, and in particular since 2010 (Lankford and Silver, 2019; Krouse and Richardson, 2015). Mass violence is consistently committed by white males in their teens to 60s, and most often in the 30s. Most of these perpetrators lack a set of pro-social factors, including high divorce rates, low education, social isolation, a recent personal catastrophic event, and “loner” tendencies (Fox and Delateur, 2014). Further, mass violence attacks are rarely spontaneous events; they are detailed and well-planned far in advance (Silver et al., 2018; Koper, 2020). Personal and individual causes of mass violence suggest a toxic masculinity exists among many offenders. Perpetrators often have a perverted sense of entitlement, a history of domestic violence, and connections to ‘incel’ subcultures associated with violence and hatred towards women (Follman, 2019). Social factors and community causes linked to mass violence include oppressive and marginalized social environments, consistent economic instability, and low civic engagement (Kwon and Cabrera, 2017). Recent mass violence appears to exhibit some degree of copycat and contagion behaviors; the 1-2 weeks following attacks are at higher risk for more mass violence (Jetter and Walker, 2018), and the prevalence of events has increased fame-seeking and offender idolizing behaviors (Silva and Greene-Colozzi, 2019; Lankford and Silver, 2019). Mass violence tends to occur in communities that are less social, less rural, have high rates of overcrowding, high incomes inequalities, greater firearm availability, and more permissive laws for firearm ownership, access, prevention, and safety (Markowiak et al., 2018; Webster et al., 2016; Reeping et al., 2019).

Thus, this study seeks to use the patterns identified here to measure attack contagion, violence, socioeconomics, mental health, and firearm preferences for communities experiencing Black Swan Shootings. The determination of these analytic layers and their subsequent social factors are grounded in the research discussed in this chapter. There are composition effects and community effects. Composition effects involve extrapolating common perpetrator behaviors to the community level. As these studies have demonstrated, prediction and threat assessment for perpetrators of mass violence using biographic and demographic traits is challenging, given the similarities to much of the general population. Thus, using those traits to identify composition effects has the potential to identify high-risk communities instead. Increases in attacks since the 2000s, coupled with copycat and modeling behaviors serve as composite effects for measuring spatial and temporal contagion effects. Offender racial and gender uniformity combined with racist and sexist ideologies serve as composite effects for measuring community heterogeneity. Social isolation and depression are common among mass shooters (Gill et al., 2016), so communities with higher rates of mental health problems and suicides may be at higher risk for attacks, as the potential offender pool is larger. And as studies such as Markowiak et al. (2018) and Kwon and Cabrera (2017) have demonstrated, communities with higher at-risk populations have been associated with higher rates of violence and mass violence. Thus, the attack contagion, socioeconomic, and mental health analytic layers are derived from composition effects identified in this literature review.

Community effects demonstrate weakened social controls and external pressures. These are the factors that potentially create an environment conducive to mass violence. This includes social disorganization factors associated with violence, socioeconomics, and access to weapons. Violent crime becomes a proxy for oppressive social environments. Population density, unemployment, and poverty measure community overcrowding and instability. Greater firearm availability and less restrictive firearm laws serve as indicators for access to weapons. Multiple of the aforementioned studies demonstrate that communities with these conditions are at higher risk for mass violence. Thus, each community-based social factor measured in this study is grounded in prior research on its relationship to mass violence.

METHODOLOGY

As previously described, a Black Swan Shooting is defined as an attack in the U.S. involving perpetrator(s) acting in a non-state-sponsored capacity using a firearm(s) to kill or injure a significantly large number of people. While this definition does not require Black Swan Shootings to be public mass shootings, most are. Similarly, this definition does not exclude terrorism; for a terrorist attack to be considered a Black Swan Shooting, a firearm must be used, and an extraordinary number of casualties caused. And finally, this definition uses firearms as a main qualifier for identifying attacks. This is central to the notion that humans may be inherently violent, but in societies where the barriers to violence and its associated weaponry are exceptionally low, the relative ease to cause mass carnage is more likely. Thus, Black Swan Shootings are an inherently U.S. problem in part because of the culture surrounding firearms.

Comprehensive identification and analysis of such events requires a five-part methodology for examining attacks and their predictability. First, there is creation of the event data, which includes the sources, the victim casualty threshold, and the temporal start point. Second, a discussion on the analytic layers used in the social factor location correlations. Third, a description of the panel time series data set using the Black Swan Shootings and social factor layers. Fourth, a detailed analysis plan. Fifth, a discussion on testing and metrics for evaluating predictions.

Expected results

This study seeks to identify the combination of social factor layers and attributes that demonstrate a spatial relationship to Black Swan Shootings over time. First, this study identifies the nature of copycat and contagion behaviors in the form of spatial hotspots and temporal sprees. Second, the initial output will be a matrix of the analytic layers examined, their degree of relationship, and potential predictability for Black Swan Shooting attack and residence locations. It will be the cumulative result of the best performing analytic layers, and the proactive application of the historical relationships between these variables and Black Swan events (if any). Third, after the logistic regression is developed and run for each year, a table of risk assessment results will be generated and analyzed. Thus, this study has the potential to be applied to the current time periods' geography and social conditions to identify the counties at high and lower risk probabilities for a Black Swan Shooting to occur during the current (or future) year. As previously mentioned, "high" and "low" risk are relative terms. The overall risk for any county in the U.S. in a given year is low, but high-risk counties identified in this model have a probability higher than normal (but still low). If there is no identified relationship between Black Swan Shootings and these analytic layers, then the findings from this study demonstrate the lack of relationship and the overall spatial unpredictability of such attacks.

Thus, this study addresses three specific questions:

1. Do Black Swan Shootings cluster in space and time? If so, when a new attack occurs, what can be inferred about the location and timing of the next attack?
2. What social factors have a relationship with Black Swan Shootings, and to what extent? This study suggests there are 11 potential factors that have some degree of relationship to these types of attacks.
3. Do social factors that have a relationship with Black Swan Shootings predict the number and location of attack and/or residence of Black Swan Shootings in a given year? And to what extent can this be modeled going forward?

Mass violence event data

Black Swan Shooting data sources

Black Swan Shooting event data is based on a combination of multiple overlapping sources. In an attempt to identify as many relevant incidents as possible and recognizing one definitive dataset of such events does not exist, a comprehensive set of records is built across 20 unique datasets. Table 1 highlights 18 of those data sources, based on mass shooting definition and collection time range.

Table 1 Black Swan Shooting data sources

Source	Mass Shootings	Victim Count Threshold	Fatalities	Casualties	Temporal Start	Temporal End	Caveat
Mass Shooting Tracker	2361	4	Y	Y	2013	2018	NA
Gun Violence Archive	1669	0	Y	Y	2014	2018	All firearm violence
K-12 School Shooting Database	1338	0	Y	Y	1970	2018	All school shootings
Brady Campaign	472	3	Y	Y	2005	2012	NA
START	417	0	Y	N	1970	2017	Terrorism events with a firearm
USA Today	361	4	Y	N	2006	2017	NA
Stanford Mass Shooting in America	335	3	Y	Y	1966	2016	Shooting must not be identifiably gang, drug, or organized crime related; minimum of 3 corroborating sources for inclusion
Congressional Research Service Report	317	4	Y	N	1999	2013	Compiled data using FBI SHR, agency press releases, and compiled events from advocacy groups
Everytown	173	4	Y	N	2009	2017	NA
School Shooters	124	0	Y	Y	1913	2018	All school shootings
Mother Jones	107	4	Y	N	1982	2018	Most significant public events
Mayors Illegal Against Guns	93	4	Y	N	2009	2013	Compiled data using FBI SHR and media reports
InfoPlease	85	0	Y	Y	1996	2018	"Most notable" culturally impactful events
Violence Policy Center	66	0	Y	Y	1980	2018	Events involving large capacity ammunition magazines
L.A. Times	47	3	Y	Y	1984	2017	"Most notable"
CNN	37	8	Y	N	1949	2018	Deadliest events
NY Crime Commission	31	4	Y	N	1984	2012	Events involving large capacity ammunition magazines
Violence Policy Center	29	0	Y	Y	2007	2015	Events involving concealed permit handgun holders

Similar to the process for defining mass murders and mass shootings, the datasets compiled in this study have multiple similarities and slight differences. The majority of the data comes from open source data compilations and aggregations of media outlets, public information releases, and advocacy groups. Seventeen of the 20 datasets are purely open source, and one of the remaining three (Congressional Research Service) is a hybrid of official data augmented with open source data. The two official datasets, both from the Federal Bureau of Investigation (FBI) are also the two most problematic: National Incident-Based Reporting Service (NIBRS) and Uniform Crime Report (UCR) Supplemental Homicide Reports (SHR) suffer from a mixture of limited agency participation, slow availability, and difficulty to query. These limitations mean NIBRS and UCR SHR are purely supplemental sources for this study, not part of the dataset construction, and thus excluded from Table 1.

Across all 20 sources, the largest discrepancy is the number of records, ranging from 2,361 (Mass Shooting Tracker, which collects all mass shootings with four or more killed or wounded victims since 2013) to 29 (Violence Policy Center, which monitored attacks involving concealed handgun permit holders from 2007-2015). However, there are two natural breaks in the event counts across these datasets. As Figure 3 demonstrates, the top three datasets (Mass Shooting Tracker, Gun Violence Archive, and K-12 School Shootings) naturally cluster with event counts greater than 1,300 events; these are followed by a group of five similar datasets (Brady Campaign, START, USA Today, Stanford, and Congressional Research Service) in the 300-500 event range; and then a group of the remaining ten datasets, each with less than 200 events (and seven of the ten having less than 100 events).

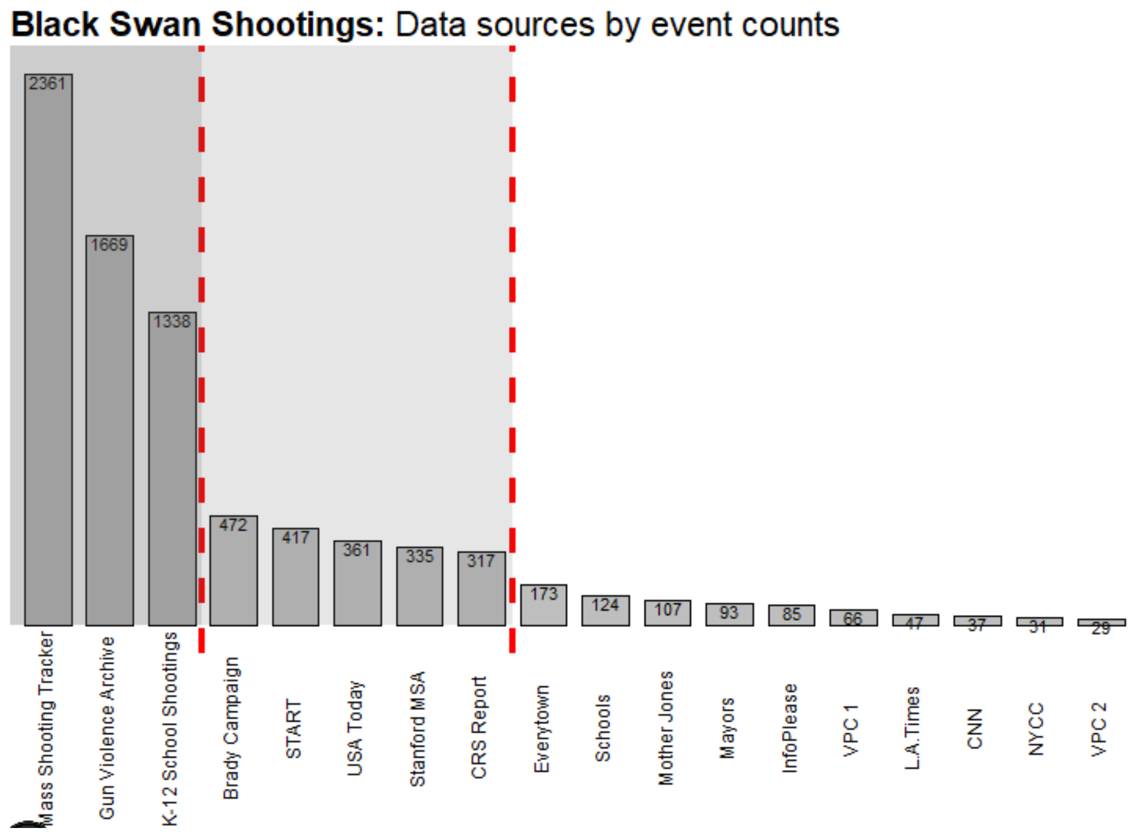


Figure 2 Black Swan Shooting sources, by natural breaks in event counts

One of the reasons for using so many datasets in this study is to build depth and breadth across Black Swan events. The larger datasets, including Mass Shooting Tracker, Gun Violence Archive, and the K-12 School Shooting Database offer a greater volume of events. The smaller datasets, including Violence Policy Center, CNN, L.A. Times, and Mother Jones offer richer details on the most significant events. These details include perpetrator characteristics and background, weapon, and location information. Regarding victim count thresholds, three datasets use a three-count cutoff, seven of the datasets use a four-count definition, one uses an eight-count threshold, and nine of the datasets do not have a clear victim count demarcation (because they are themed-based, such as exclusively school shootings or terrorism). All of the datasets track fatalities, but two of them (NIBRS, UCR SHR) are exceedingly difficult to find information on multiple victims. Half of the datasets include wounded victims as part of their casualty count thresholds. While some of the datasets go back to the early 1900s, most of them are relatively new; 15 of the 20 datasets have event start points in the 1980s. Eight of the datasets begin in the 2000s, two in the 1990s, and five start compiling events in the 1980s. Similarly, 15 of the 20 sources end in 2016 or later; eight are actively compiled today, five are built annually (thus currently cut-off at 2018), and two ends in 2016. It is worth noting that two datasets, Gun Violence Archive and Mass Shooting Tracker, have the shortest duration of collection (six and seven years, respectively), but the largest record sets, and provide typically 24-hour or less updates.

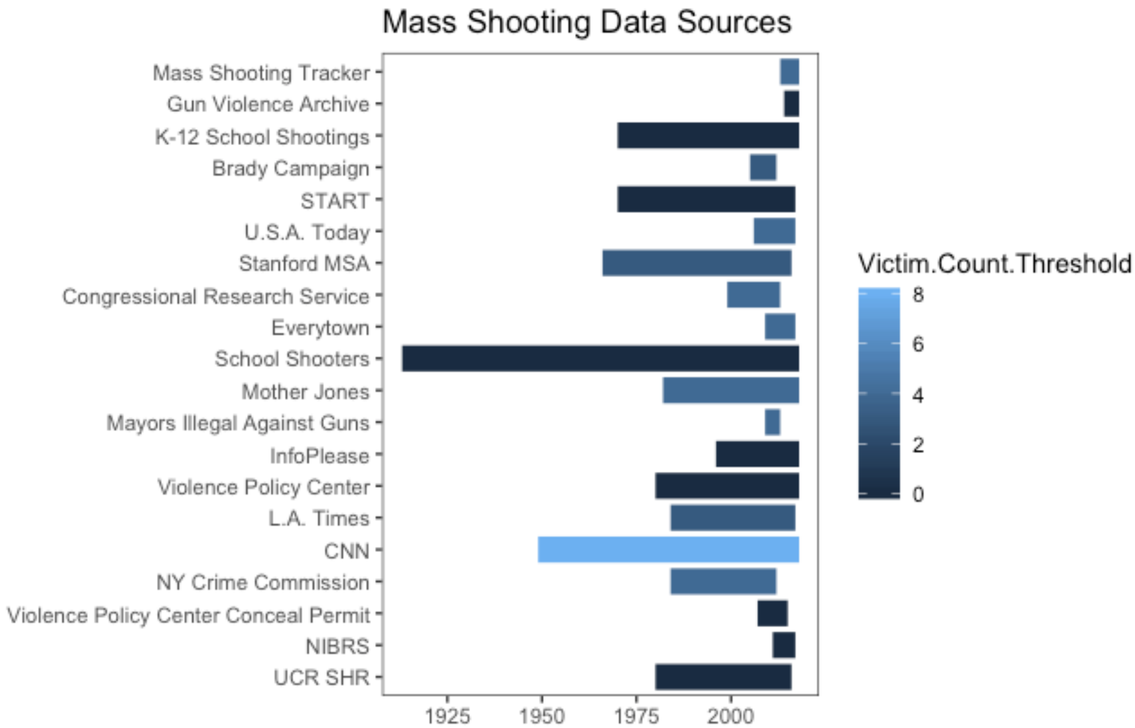


Figure 3 Black Swan Shooting sources, by collection duration and victim thresholds

From a purely aggregated perspective (and prior to any deconfliction and de-duplication), the initial dataset compiled across all sources would consist of events of varying casualty counts since 1913. Since the attributes across this corpus are not consistent, this approach will not work. As mentioned before, the longest spanning datasets are highly specialized, and capture distinct event types, such as school shootings or terrorist attacks without a specific victim count threshold, or arbitrarily defined "notable" attacks. With such uneven collection mechanisms, aggregating and compiling events is problematic, as these niche datasets are best leveraged for enriching the event details. Simple statistical breakpoints in count and time cannot be identified when merged attributes and timelines are inconsistent. These initial 20 sources categorize into official sources, open/unofficial sources, and mixed sources. Each group has clear advantages and disadvantages.

Both official data sources originate from the FBI: the UCR SHR, and NIBRS. For both sources, data is initially recorded and aggregated by originating state and local law enforcement jurisdiction and provided to the FBI on a scheduled basis (Lynch and Jarvis, 2008). The advantages of using these data are validity and a relative degree of accuracy. UCR data has been collected since 1929, and while voluntary, it has long been considered the most comprehensive and consistent crime data across the country (Huff-Corzine et al., 2014). NIBRS focuses on incident-level details, including case clearance and a perpetrator's cause of death (Huff-Corzine et al., 2014; Addington, 2004). Fegadel and Heide (2016) compared homicide arrests and found despite NIBRS encompassing 30% of the U.S. population, data for victim, offender, and weapons are consistent between NIBRS and SHR. Regarding accuracy, Huff-Corzine et al. (2014) note that law enforcement agencies are typically reluctant to share investigative data before a case is closed. Thus, when an offense is eventually entered into SHR it will lack errors and have greater validity than media reporting.

Disadvantages with using official data sources include the breadth, speed, and accuracy of coverage. Duwe (2000) studied mass shootings by comparing SHR to detailed news accounts, and found that the SHR slightly underreports homicides, uses vague coding, and contains missing and erroneous variables (including offense location). Further, research indicates NIBRS covers approximately 9% of law enforcement agencies, and Huff-Corzine et al. (2014) note historical studies using NIBRS undercount mass murders annually. UCR has similar gaps in space (missing data at the county level) and time (missing data for entire years), with inconsistent methods for imputing the gaps (Maltz and Targonski, 2002; Lynch and Jarvis, 2008). Multiple studies found that despite capturing 90% of all UCR homicides, SHR inconsistently omits events (including the 2012 Sandy Hook attack), erroneously aggregates homicide data (including weapon type) for some years, mislabels victim-offender relationships for multiple victim events, and lacks details on weapons (Fox and Levin, 2015; Huff-Corzine et al., 2014; Addington, 2004). Overberg et al. (2013) compared open source mass shooting events from 2006-2011 to SHR and found just 61% of the events in FBI holdings. Maltz and Targonski (2002) describe the SHR data collection as "inconsistently inconsistent." Thus, Huff-Corzine et al. (2014) recommend using SHR for aggregated annual mass murder counts, as essential information (such as victim counts) are unreliable. UCR and NIBRS reporting are typically conducted in monthly increments. The F.B.I. makes the data publicly available after ingesting and aggregating from across the country. Any given month's crime data may not be available until two months later (best case scenario), or upwards of a year or more (often more).

One final note on official data source disadvantages: another source of potential value exists - the National Crime Victimization Survey (NCVS). Since the 1970s, twice a year more than 40,000 households are interviewed for information about crime (F.B.I., 2018). The NCVS is a sample with which to derive trends; it is not a comprehensive set of records to query specific events for investigation. Thus, the N.C.V.S. is unable to be used in the current study.

Using open source event data also offers advantages and disadvantages. Advantages include longevity of collections; richness of details; and speed. Unlike SHR, which was first available in 1976, open source media has been around in different formats for centuries (Huff-Corzine et al., 2014). Second, despite lacking law enforcement investigative details not shared with the public, the level of detail available via open source data is often richer and fuller. Huff-Corzine et al. (2014) found that variables such as offender and victim names, history/backgrounds, exact locations of occurrence, law enforcement response, and a thorough timeline of events are typically best derived from open source.

Third, open source data is incredibly timely. In the age of heightened awareness and omnipresent news cycles, events are reported globally in near real-time. Within 24 hours or less of occurrence, most of the particulars surrounding an attack are made public. More recent studies of mass shootings (Dillon, 2013; Huff-Corzine et al., 2014) have favored using open source data over official data. The most notable disadvantage to using open source data for mass shooting events is inaccuracy. This is the natural result of speed, conflicting reports, and the economics of media reporting. Duwe (2007) found media coverage has a greater tendency to contain errors in the immediate aftermath, and also tends to distort views of the offenders and victims. Further, what qualifies as newsworthy is not uniform, as Duwe (2000) noted that nearly all mass murders are locally important, but what resonates at a national scale can be difficult to anticipate (given competing news priorities, among other variables). This makes data acquisition more difficult, having to span multiple diverse sources to ensure broad coverage. Thus, open source collection has occasionally led to studies narrowly focusing on the widely recognizable cases, and neglecting many of the less noteworthy, non-public, and familial events (Duwe, 2000; Duwe, 2007; Huff-Corzine et al., 2014). In the immediate aftermath of a significant catastrophic event, value in the open source is placed on speed rather than correctness. This is not entirely problematic, as the information is updated as time passes. But initial data has a greater tendency to have erroneous elements as well as contradictory reporting. Further, “unreported” mass shootings happen much less frequently; the rise of several online data sources aggregate crimes on a daily basis, including Mass Shooting Tracker and Gun Violence Archive.

A third option is a mixed methods approach, that combines official and open source datasets. Such an approach potentially increases the validity, by using the best qualities of both types: the consistency from official sources, and the richness, availability, and speed from publicly available sources. NIBRS has the case details, but not the speed and geographic coverage. SHR has nationwide coverage, but not the details such as event location and weapon types. Open source has the speed and some of the missing data elements. This is not a new concept, as Huff-Corzine et al. (2014) found that many prior trend studies of mass murder used SHR, but supplement their data sets with open source media reporting, including newspapers, television news, and weekly magazines (Petee et al., 1997; Duwe, 2007; Fox and Levin, 2015; Fox and Levin, 1998; Dillon, 2013).

The main negatives to a mixed data type approach are administrative processing and timing. Administrative processing refers to identifying, cleaning, and merging redundant records across multiple disparate datasets. A non-trivial amount of data wrangling is involved in creating (and maintaining!) such a data set, as each source has similar yet unique ways to capture data. For example, different definitions of mass violence lead to differences in collection. Fusing data means significant cleaning and normalization. Timing relates to uneven timelines of production. Open source data is typically available within hours or days after an event occurs, while UCR and NIBRS datasets can take months or even years to populate. However, all of these are secondary issues compared to data quality and completeness - the two biggest issues that merging official and open sources solves.

Given these pros and cons, the data set for this study will be mixed, using open source data as the foundation of events, and leveraging official data sources to confirm, amplify, and enrich those attacks. This study will be relatively unique in that it leverages open source, online databases that regularly and systematically scrape hundreds of news mediums and media platforms. This enhances the ability to quickly update findings in the future, recreate this study, and even update/modify definitions for different versions of Black Swan attacks. The primary open source datasets give this study a real-time component for future operational utility. However, given the definition of Black Swan Shootings and these available data sources, there remains several components that need to be further developed: the victim count cutoffs, the temporal start, and how to classify victims/casualties. As such, determining these thresholds will be done with a data-driven approach.

Data subset to determine events

Due to definition and data collection inconsistencies, a subset of mass shooting sources will be used to identify the remaining components of the Black Swan Shooting definition. A subset of the most recent and most comprehensive data is used to build a baseline of normal mass shooting activity. Although this sample has an inherent recency bias, with advances in technology, media, and open source collaboration, this method ensures the most complete picture. Using an appropriate subset will allow for proper thresholds to be created, and then reapplied to all 20 datasets to find matching events. This process *uses* the data to *find* the relevant data.

The data subset for identifying victim count thresholds will be based on a combination of the Mass Shooting Tracker (MST) and the Gun Violence Archive (GVA). It is worth using both MST and GVA for several reasons. First, both sources have events that the other does not. Second, merging them can aggregate details: over time, both have become relatively consistent for victim counts and dates, but GVA often provides specific addresses, and MST provides victim names and links to additional details. Thus, merging the two enriches the content and confirms maximum coverage during the subset time period, 2013-2018. MST captures more events (albeit marginally so) and GVA captures more details.

Despite its high volume, the K-12 School Shooting database is excluded from this subset because it is all school shootings, not just mass shootings at schools. Since the goal is to truly identify the most prolific shootings within the largest amount of noise (so, create the largest sample size of similar event criteria over the smallest possible time period possible to find true outlier casualty counts), adding extra data sets that cover a specific type of event over longer time periods would artificially skew the significance by ignoring all the non-school shootings that occurred prior to the MST and GVA collections.

Thus, combining the two datasets results in all U.S. mass shootings involving four or more killed or wounded victims, from 2013-2018; a combined 4,037 initial records. This set is then de-duplicated by date, city, and victim count. This process accounts for reporting artifacts, including 1-day lags, city vs. county location descriptions, and 1-2 victim count mismatches. When matching records differ by one day, the earlier record is chosen. When matching records differ in victim counts, the larger quantity is chosen. After applying these filters, the dataset is reduced to 2,448 mass shootings during 2013-2018. Using this subset of mass shooting events, a victim threshold for casualty counts (killed and wounded) can be determined. Then, that cutoff can be applied to all of the mass shooting datasets collected to identify a natural, self-organizing temporal start. These victim and temporal thresholds are by no means arbitrary; rather, they are based on natural breaks and anomalies in the collected data.

Victim threshold

There are multiple methods to determine a victim cutoff. Realizing that Black Swan Shootings are intended to be the rarest of the rare, an appropriate victim count should capture the highest event casualty counts and remove the noise of “ordinary” mass shootings. One method to select a cutoff is to choose one that will maximize the number of cases defined as mass murder (so there are enough events to study) while minimizing the likelihood that the cases defined are murders occurring as a byproduct of other crimes, which would otherwise confound research findings. The Bureau of Justice Statistics (BJS) publishes annual survey data on criminal victimization other than homicide, including the number of victims per incident. For example, for all violent criminal incidents in 1983, 88.5% involved one victim, 8.9% two victims, 1.7% three victims and 1.0% four or more victims (Dietz, 1986). A threshold for Black Swans of three victims would exclude the possible byproducts of more than 95% of violent crimes, and a threshold of five victims would exclude 99% of violent crimes. Such a method would suggest five fatalities as the threshold for Black Swan Shootings. However, as Figure 4 illustrates, using the aforementioned data subset, such a threshold would include at least 81 attacks during 2013-2018, and subsequently hundreds of shootings for a larger time range. Thus, this threshold would be capturing too much "ordinary" mass shooting activity.

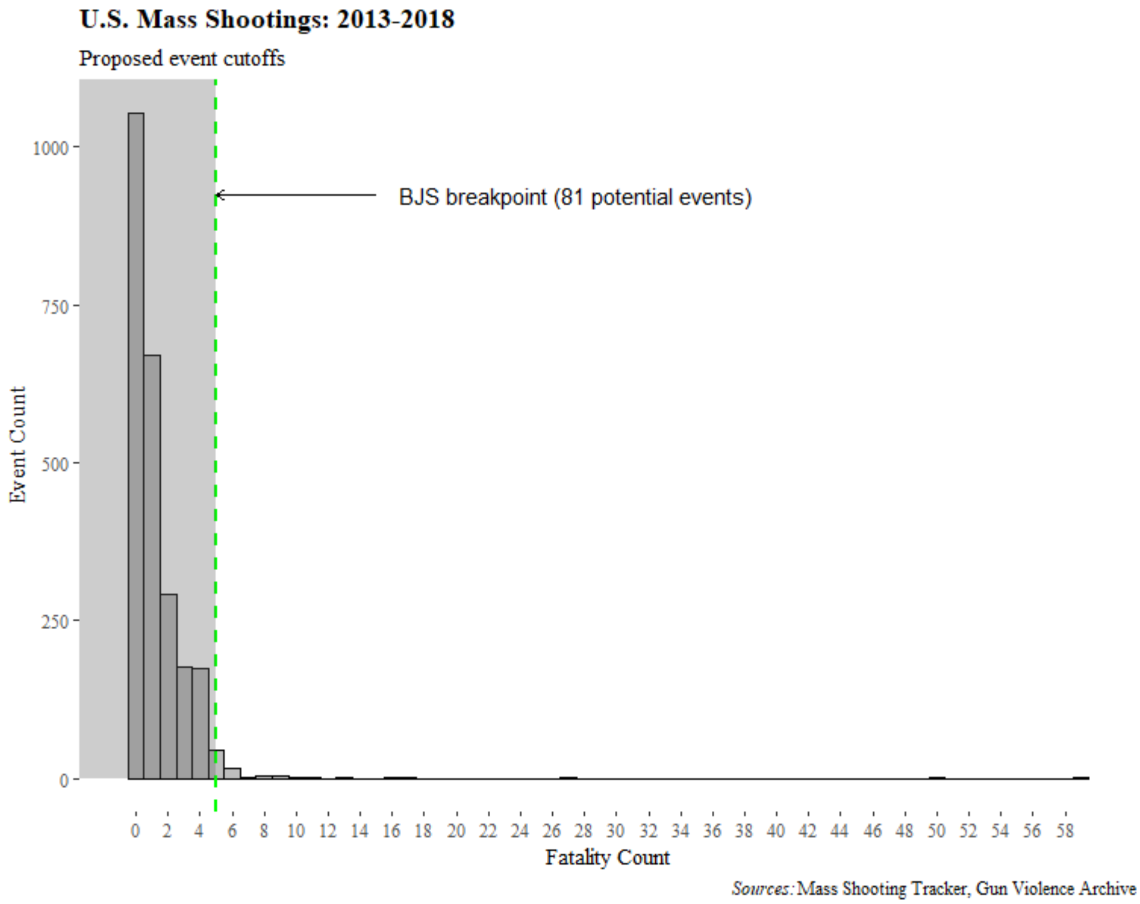


Figure 4 Mass shooting data subset total victim counts, with BJS breakpoint

Another method for defining a Black Swan Shooting victim cutoff would be academic sources. Prior research for mass murder offers multiple victim count cutoffs, with most focusing on 3-4 people killed per event. The United States Congress, as well as multiple academic studies, define mass murders as three of more people killed in a single incident (Dowden, 2005; Meloy et al., 2004; Meloy et al., 2001; Hempel & Richards, 1999; Levin & Madfis, 2009; Holmes & Holmes, 1992; Petee et al., 1997; Ho Shon & Roberts, 2010; Allely et al., 2014; Krouse & Richardson, 2015). A three-victim fatality count is often used because it logically represents the lowest count that eliminates single- and double-homicides (Levin & Madfis, 2009). Three fatalities also exclude one victim-plus-suicide events and categorizes everything else as a mass murder. However, the FBI, along with many prevalent research studies, determine mass murders to involve four or more people killed (Fox & Levin, 1998; Duwe, 2004; Lankford, 2015; Hilal et al., 2014; Carcach et al., 2002; Duwe, 2000; Fox & Levin, 2003; Knoll, 2010; Knoll, 2012; Duwe, 2005; Lankford 2015b; Huff-Corzine et al., 2014; Dillon, 2013). Four or more killed is often cited as a common threshold for mass murder because it clearly differentiates from double homicides (Fox & Levin, 2003), and it is less prone to measurement errors, compared to a three-fatality cutoff (Duwe, 2004). One contemporary study defines mass murder as something greater: Lester (2002) uses five or more killed as the threshold.

Research on mass shootings has a wider definitional variance for victim count thresholds. Typically, mass shooting academic research uses four or more killed as its threshold (Dahmen et al., 2017; Jeltsen, 2015; Cohen et al., 2014; Team Trace, 2017; Fox & Levin, 2015; Fox, 2013; Everytown, 2020; Krouse & Richardson, 2015; Lankford 2016; McPhedran & Baker, 2011; Dillon, 2013; Cannon, 2016; Follman et al., 2018; Lankford 2016b; Bjelopera et al., 2013). Several media outlets and publicly available open sources have also used four or more killed or injured as their reporting threshold (LaFraniere et al., 2015 [New York Times]; Guo, 2015 [Washington Post]). Both of these criteria are in line with mass murder research thresholds. Other studies have focused on two or more killed or injured, a threshold much lower than mass murder research (Lankford, 2015b; Lankford, 2012; Newman and Fox, 2009). Such a cutoff is designed as everything greater than one, as multiple-victim incidents are seen as an entirely different phenomenon than single-victim murders (Lankford, 2015b). In some cases, six or more killed (Chapman et al., 2006), and six or more killed or injured (Kleck, 2016) have been used. In one study, Kleck (2016) argues the six-victim threshold is functional, as there are six bullets in the traditional revolver, and any shooting causing more than six casualties requires the offender to reload. Obviously, this does not take into account assault weapons, large capacity magazines, or the ability of a single bullet to inflict multiple casualties, but the approach is noteworthy. Inevitably, as Kleck (2016) argues, whatever the cutoff is, it will be somewhat arbitrary. Also, these cutoffs are baselines for mass murder and shootings. This study is looking to find the extreme versions of these attacks. If the BJS cutoff was too low and encompassed too many “ordinary” mass shootings, these lower thresholds, as evidenced by Figure 5, are certainly no better.

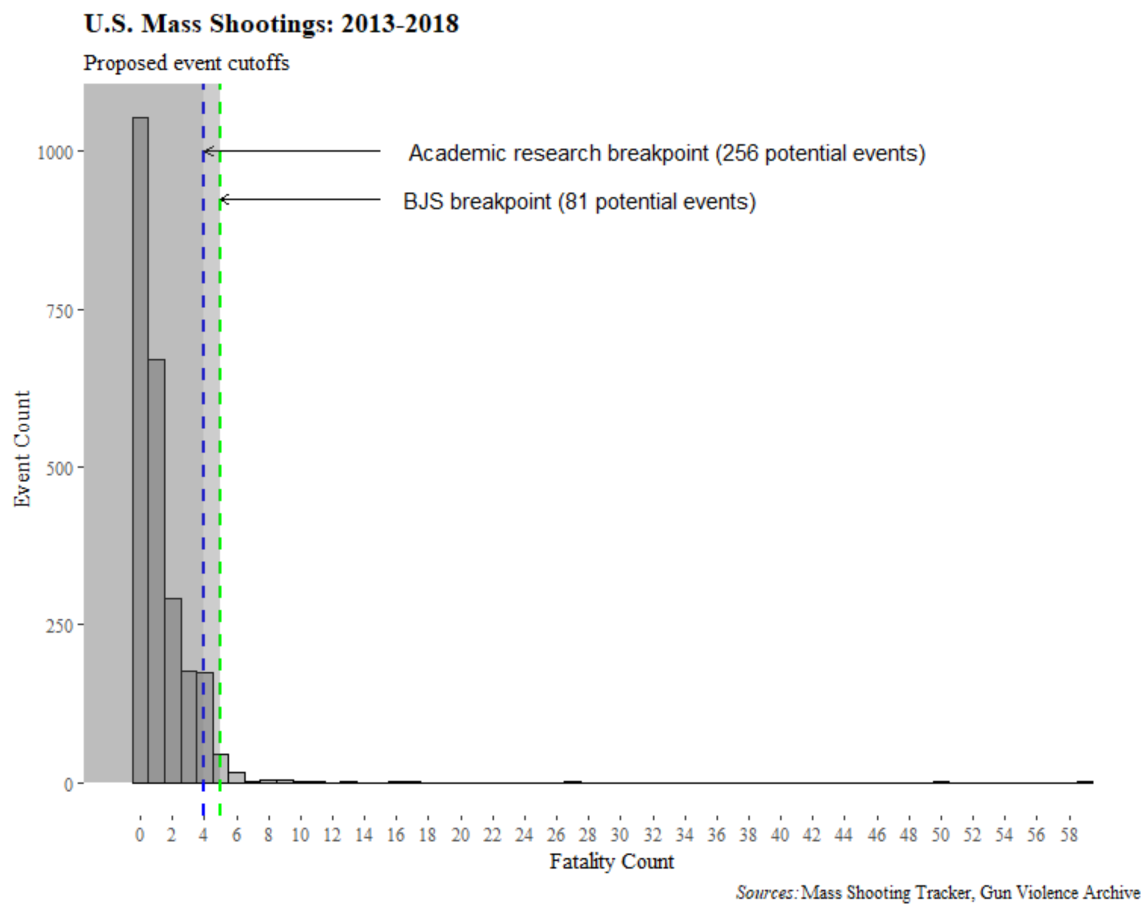


Figure 5 Mass shooting data subset total victim counts, with BJS and academic breakpoints

The preference of this study is to let the data self-organize, and determine a threshold based on a significant, naturally anomalous significant breakpoint in the event data. Thus, this ambiguity among previous research does not clearly inform on a victim threshold for Black Swan Shootings. Any cutoff will need to be significantly higher than these. This research employs a three-pronged analytic method for determining victim count cutoffs: measures of central tendency; sequential percent changes; and anomaly detection. These similar but slightly different methods are used to identify natural breaks in the data. Each of these methods is calculated for counts of killed, wounded, and total casualties. The method(s) determined to be the most rigorous will be applied to the master data set, keeping records that match any of the three casualty types. The resulting list will then be subject to a qualitative review. Ultimately, the event data in this study should be derived from a mix of empirical, justifiable methods, as well as align with contemporary theories on mass violence. Thus, each event included in the initial Black Swan Shooting dataset will be critically analyzed to ensure it aligns with empirical and theoretical standards.

Measures of central tendency include calculating the mean, median, mode, skewness, and standard deviation. This set of statistics are used to develop a sense of normal/average activity. Once established, detection of outliers and extremes can be based on standard deviations from the mean. To determine the type of data distribution, the skewness of the event casualties can be measured. In such a model, event casualty counts exceeding the mean plus two standard deviations would fall outside the upper bounds of 95% of all cases, and could reasonably be assumed to be distinct outliers, and considered Black Swan Shootings.

Table 2 Mass shooting data subset measures of central tendency

Type	Mean	Median	Mode	Skew	Standard.Deviation	Max	Min
Killed	1.27	1	0	12.25	2.26	59	0
Wounded	3.86	4	4	44.57	9.16	441	0
Total Casualties	5.13	4	4	43.99	10.44	500	4

The breakpoints here would be events with either 3 killed, 13 wounded, or 15 total casualties. These results indicate that the wounded and total thresholds appear sufficiently high; but the killed count is low and would incorporate far too many events. Confidence in this model is relatively low, given the extreme skewness of the data set; as Table 2 demonstrates, the skewness for total casualties is 43.99. Even though this is not the final answer, it demonstrates the magnitude of the cutoff, and how this study differs from traditional mass shooting research.

The second method for analyzing a victim threshold is examining the sequential percent change for each bin of casualty counts. This method is designed to work better with skewed distributions that exhibit non-linear/non-parametric changes and detecting the specific pairs of counts where substantial change occurs. Although typically applied to temporal data, as described by Paulsen et al. (2010) when analyzing crime rates, this study is able to bin and sequence the number of killed, wounded, and total casualties and analyze in a similar manner.

Table 3 Mass shooting data subset casualty counts and percentage changes

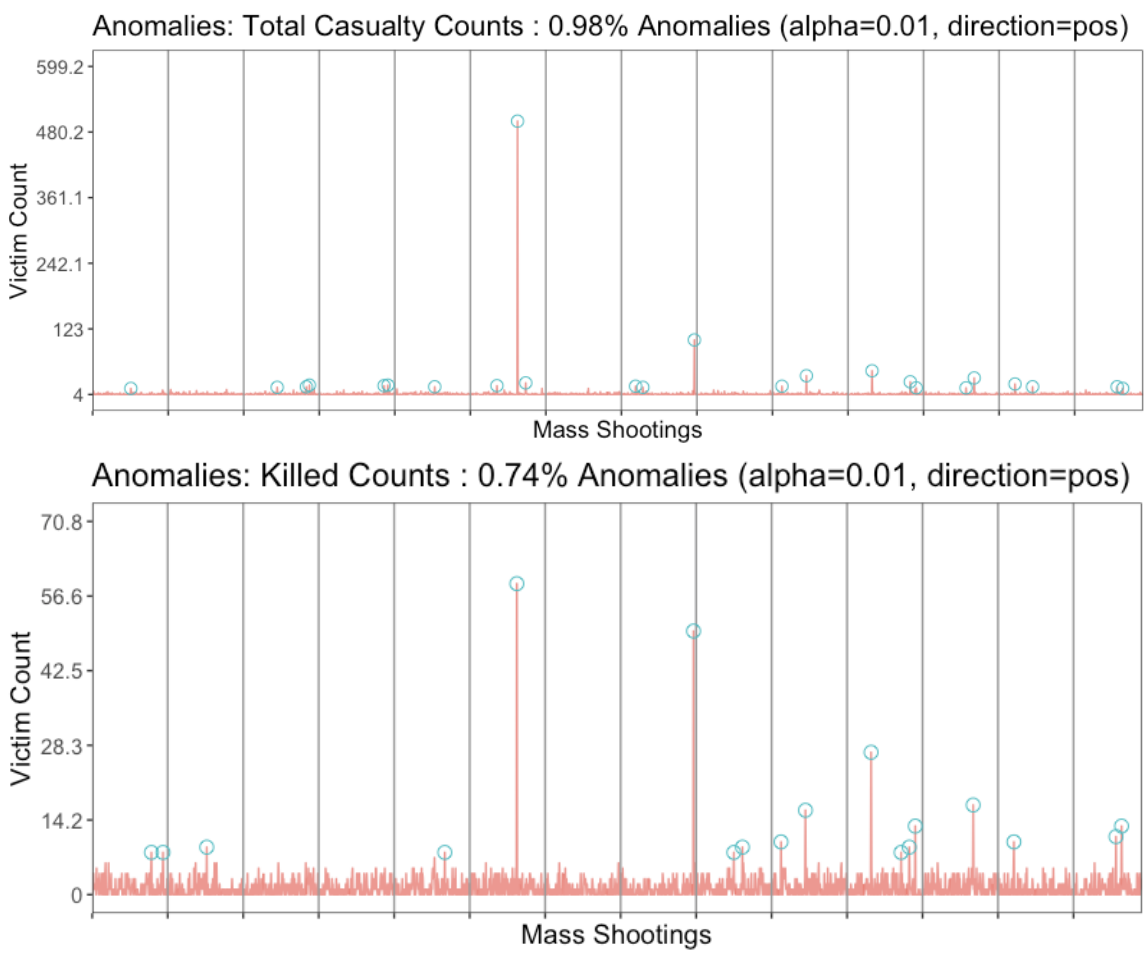
Total Casualties	Frequency	Percent Change
4	1477	
5	555	-62.4
6	178	-67.9
7	111	-37.6
8	51	-54.1
9	21	-58.8
10	12	-42.9
11	4	-66.7
12	8	100.0
13	3	-62.5
14	2	-33.3
15	4	100.0
16	2	-50.0
17	2	0.0
18	4	100.0
19	2	-50.0
20	2	0.0
21	2	0.0
23	1	-50.0
25	1	0.0
27	1	0.0
34	1	0.0
38	1	0.0
47	1	0.0
103	1	0.0
500	1	0.0

Fatalities	Frequency	Percent Change
0	1053	
1	671	-36.3
2	291	-56.6
3	177	-39.2
4	175	-1.1
5	46	-73.7
6	16	-65.2
7	1	-93.8
8	5	400.0
9	3	-40.0
10	2	-33.3
11	1	-50.0
13	2	100.0
16	1	-50.0
17	1	0.0
27	1	0.0
50	1	0.0
59	1	0.0

Wounded	Frequency	Percent Change
0	179	
1	137	-23.5
2	240	75.2
3	472	96.7
4	858	81.8
5	306	-64.3
6	114	-62.7
7	69	-39.5
8	23	-66.7
9	13	-43.5
10	7	-46.2
11	4	-42.9
12	2	-50.0
13	5	150.0
14	3	-40.0
15	2	-33.3
16	2	0.0
17	3	50.0
18	1	-66.7
19	3	200.0
20	1	-66.7
22	1	0.0
25	1	0.0
53	1	0.0
441	1	0.0

As evidenced in Table 3, 99% of events have 14 or less total casualties, but at 12 total victims is the earliest, most drastic sequential percent change. Events with five or fewer fatalities comprise 99% of events. However, the first natural break in fatality counts is at 8. For wounded counts, 99% of events have 11 or fewer injured. The natural break here is at 13. In each case, after that initial difference in percent changes, the trend remains relatively similar and flat. For total victims, there are several drastic changes at 15 and 18 victim counts, but these involve relatively low raw numbers. For killed and wounded counts, each has one additional drastic percentage change, but otherwise levels out. This pattern helps confirm the notion that the percent change is a significant break in the data. Thus, this method would put the victim threshold at events with either 8 killed, 13 wounded, or 12 total victims. While there is slightly more confidence in this model, it is not empirically robust and lacks a strong theoretical underpinning of prior use in such studies. However, it does confirm some of the measures previously identified for wounded and total victim counts and raises the killed count threshold. Thus, in conjunction with the first method, this process helps build a more realistic threshold for this study.

A third model is anomaly detection, which is the systematic process of identifying abnormal, outlier elements in a dataset that fall outside of typical or expected behavior (Kerjariwal, 2015). While anomaly detection is traditionally applied to temporal datasets, much like the sequential percent change analysis, this dataset can be run for anomalies using ordered bins of killed, wounded, and total casualty counts. Often, obvious anomalies are identified along distinct temporal changes, such as yearly and seasonal patterns. Known as global anomalies, these patterns are the above or below average outliers. The simplest calculation for global anomalies is identifying events outside of two standard deviations (Kerjariwal, 2015). This would be no different than the first method, using measures of central tendency. However, a more refined process is identifying local anomalies, using the Seasonal Hybrid Extreme Studentized Deviate (S-H-ESD) model, which accounts for cyclical decomposition and piecewise approximation in non-temporal and unordered data sets (Smith, D., 2015; Kerjariwal, 2015; Rosner, 1983). Thus, S-H-ESD is perfect for detecting anomalous victim counts among a set of mass shootings. For this study, the model is calibrated to 0.01, meaning it is designed to identify the outliers that are 99.99% different than the rest of the dataset. This calibration was chosen over an arbitrary 5% or 10% because it represents approximately three standard deviations in normally distributed data (which this dataset is not), and it is the closest value to zero, so it is designed to represent the rarest of rare events. Thus, S-H-ESD can identify the most significant outliers and the truly least ordinary mass shooting casualty counts.



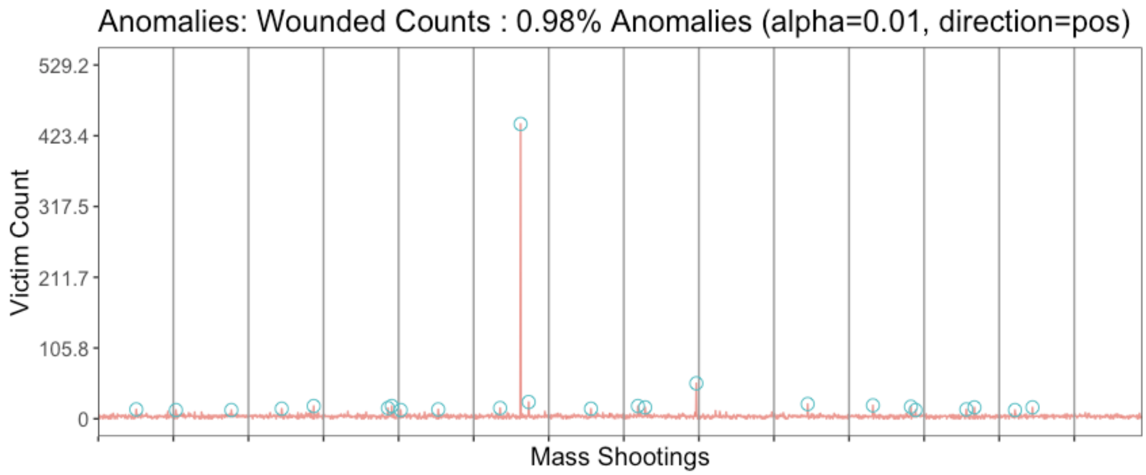


Figure 6 Mass shooting data subset anomaly detection for total, killed, and wounded victim counts

By examining local outliers and anomalies in the combined MST and GVA data set using the S-H-ESD model at the 0.01 significance level, the breakpoint is determined to be events with either 8 killed, 13 wounded, or 15 total casualties. Among the three methods, anomaly detection represents the most statistically robust (subject to calibration and adjustments), theoretically sound (grounded in skewed, non-temporal use cases), and analytically applicable process. The first two methods inform on the trending and patterns of victim counts and demonstrate the general range of anomalous events. As evidenced by Figure 8, the third method refines those ranges and creates a usable cutoff.

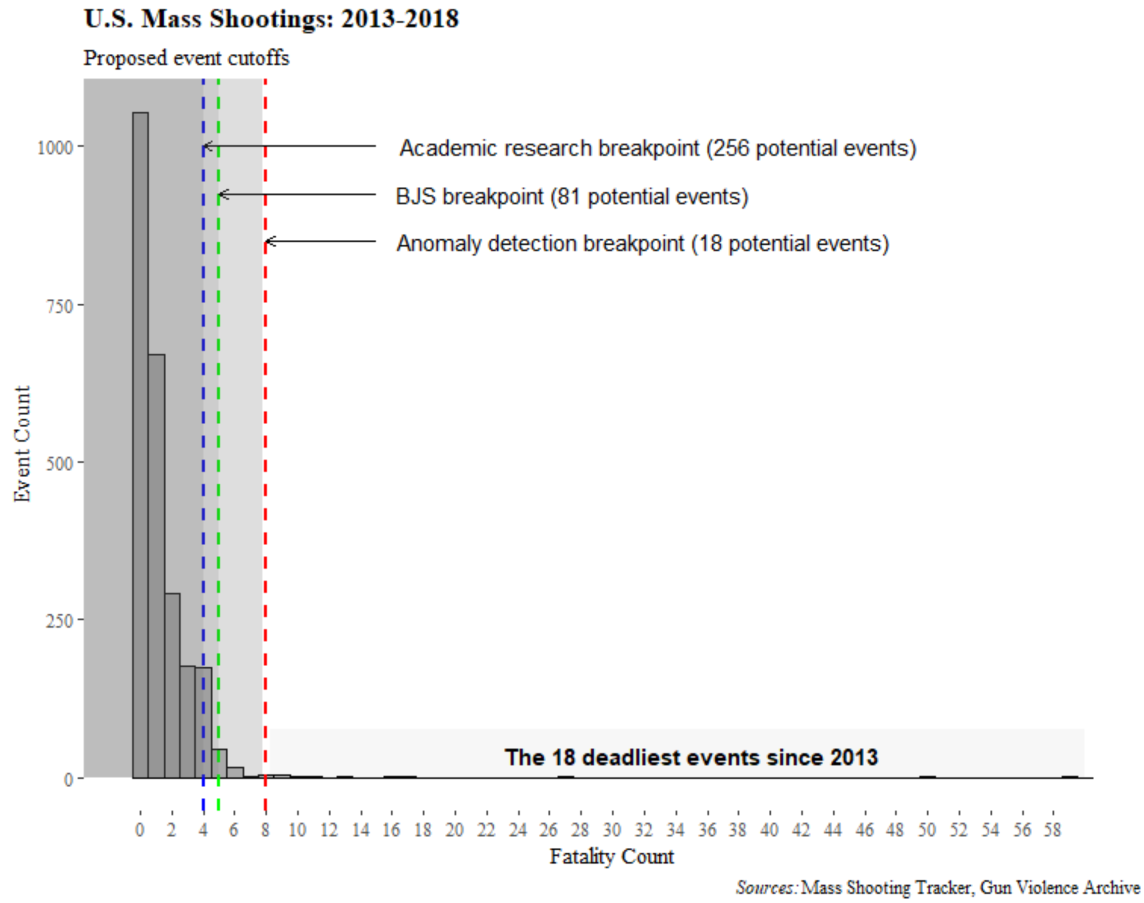


Figure 7 Mass shooting data subset threshold cutoffs for BJS, academic, and anomaly detection methods

Thus, these thresholds derived from the data subset can be applied to the original 20 data sources to identify Black Swan Shootings. However, 18 of the 20 datasets are used to merge. Two datasets - the CRS, and School Shooters - were not included. As previously described, the CRS 'data' is a compilation of annual counts of mass shootings based on other datasets already used in this study. The School Shooters dataset did not provide victim counts. Two other datasets are not used here and leveraged for event enrichment later. The FBI NIBRS and UCR SHR, due to the aforementioned timeliness and query issues, will provide supplemental qualitative details for the qualifying events. The merging and de-duplicating of the remaining 16 datasets based on the three measures of anomaly detection for victim counts identifies 86 potential Black Swan Shooting events, occurring 1949-2018.

However, a purely quantitative model for identifying events leaves some questionable results. This process overvalued low-fatality, high-wounded attacks. Such events, while severe and dangerous, are not in the spirit of a "black swan" event. In fact, such attacks fall into a unique subset of attacks involving multiple unknown perpetrators attacking each other in public places, often during celebratory events. Such attacks have occurred at street fairs, bars and nightclubs, and house parties, among others. A key differentiator here is the mutual attack from two opposing people or groups. These events do not align with much of the prior research on mass violence and mass shootings. Thus, despite meeting the statistical threshold, these events do not inform on the topic of Black Swan Shootings and modeling behavior. Thus, a qualitative review of the 86 events reduces the dataset to 62 attacks.

Temporal cutoff

There two reasons for identifying when to start the Black Swan Shooting dataset. First is data collection. Event and social factor data need to be accurate and consistent and cover a long enough time period to analyze relationships as comprehensively as possible. Some events match the casualty thresholds from the 1940s and 1960s; however, the social factor data does not go back as far. Second is relevance. While events from the 1940s or 1960s would bolster the event count, it is questionable how much they would inform on contemporary activity. Given the dynamic nature of social trends over time, a location-based social factors analysis such as this needs to capture actionable, applicable insights to the current environment. There are three different ways to potentially determine a temporal start for this study. First, there is examining and applying previously established standards. Second, there is analyzing the data and identifying a natural, self-organizing start. Third, there is combining the two first methods to identify the most logical, defensible start.

Among prior standards for studying mass violence, there is previous academic research, contemporary cultural cues, and legislative correlations. Unlike victim count thresholds, temporal start points are less consistently defined in academic research. Mass murder and mass shooting studies seem to favor going as far back as the data permits. Studies have used different specific starting points, including: 1900 (Duwe, 2004); 1966 (Lankford, 2016b); 1984 (Cannon, 2016; Follman et al., 2018); and 1994 (Kleck, 2016). One of the dangers in this study of using the earliest start point possible, is that data for location-based social factors do not go back to the 'first' Black Swan Shooting (1949), making modeling exceedingly difficult. Other studies have opted for more dynamic starts, citing notions of national attention or prominence, and using the most recent five years of data (Jeltsen, 2015). This is similar to using a patient zero approach (albeit more arbitrary) or basing the start on the earliest major contemporary event.

Second, aside from academic research, there are other society markers to base a start on. Similar to Jeltsen (2015), these include contemporary/pop culture/technological cues. Such a method would encompass the rise of the Internet in the late 1990s, which has led to an increase of information sharing and more comprehensive reporting. Third, another method would involve using a defining legislative action directly related to mass murders and mass shootings as a start point. This could include 1994 or 2004, which mark the issuance and subsequent expiration of the federal assault weapons ban. Theoretically, all attacks after 1994 involve reduced access to some (not all) high capacity magazine weapons. Consolidating these different approaches would steer the starting point of this study somewhere in the range of the mid-to-late 1990s.

Aside from these methods, an analysis of the all the qualifying Black Swan Shootings can identify a natural, self-organizing start point. This would be similar to the methods employed in developing victim cutoffs. However, while it would seem logical to apply the anomaly detection used earlier, these temporal data do not lend themselves to that model. Since this is a study of rare events, there are extremely low counts (many early years have zero or one event). Thus, this study can apply event per year and interval analysis to these events and identify the largest gaps in activity. The latter year of the largest, most recent gap can determine the start point for events in this study.

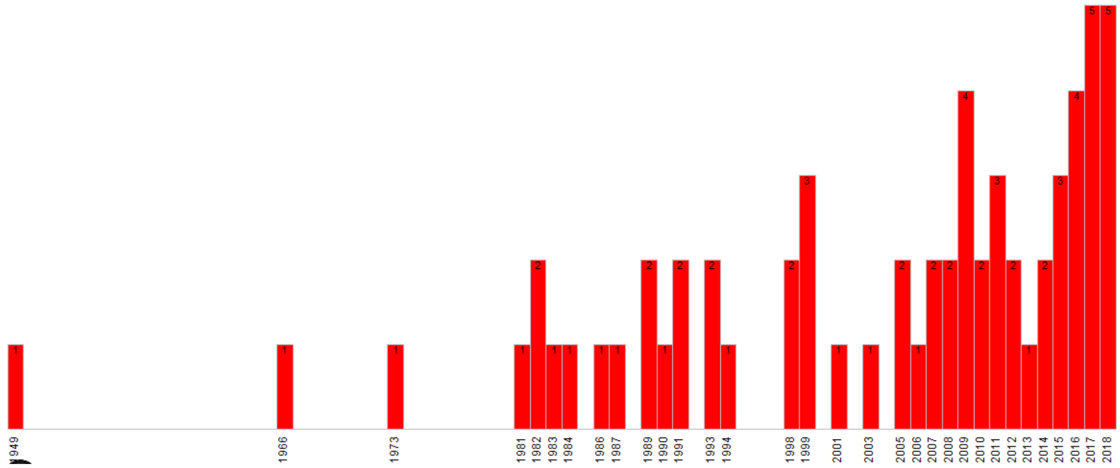


Figure 8 Black Swan Shootings per year, 1949-2018

As evidenced from Figure 8, attacks have steadily increased since the 1980s, and in particular during the mid to late 2000s. During the 1940s-1970s, there were three attacks total. During the 1980s, there was roughly one attack per year, with six of those years having 0-1 attacks. During the 1990s, there were 11 attacks, with four years having two or more, and four years having none. Noticeably, there is a three-year lull of attacks during 1995-1997; this is the longest stretch of no activity since prior to 1981. Also, this decade featured the first year with three attacks (1999). During the 2000s there were 13 attacks, including two years with no activity. During the 2010s, there have been 27 attacks, with every year having at least one attack. Thus, the window during the mid-1990s represents the largest contemporary gap and makes 1998 a logical potential starting point.

Second, an interval analysis can be conducted. This method examines the number of days between Black Swan Shootings. Similar to the analysis from Follman et al.

(2018), this process can identify significant irregular gaps in attacks over time at a more precise level. Using this analytic to identify a start point is similar to the percent change process applied to the victim counts; the goal is to find a natural break/change, followed by a leveling-off/normalizing of activity over time.

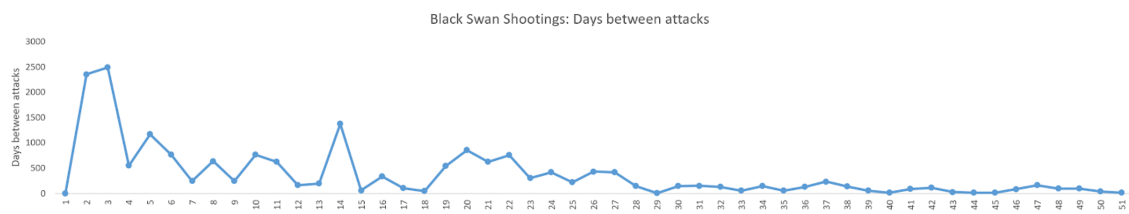


Figure 9 Black Swan Shooting intervals between attacks, 1949-2018

As Figure 9 demonstrates, the leveling off of intervals begins around attack 14 (Jonesboro, Arkansas on 3/24/1998). "Leveling off" can best be understood as a smoothing of the days between attacks into a regular and consistent timing. The gap between Jonesboro and the prior attack, 6/20/1994 at Fairchild Air Force Base, Washington, represents the largest contemporary interval (1,373 days). This is also 392 days before the Columbine High School attack on 4/20/1999. Since 1998, there has not been a clear break in the attack tempo. The next interval potentially close is between 7/8/2003 (Meridian, Mississippi) and 3/21/2005 (Red Lake Indian Reservation, Minnesota). However, this 855-day interval is still 518 days shorter than the Jonesboro-Fairchild interval. The other potential starting point to examine would be when the tempo of events has flat-lined. Immediately after the Dallas, Texas attack (7/7/2016) there have not been attacks more than 324 days apart. However, using this event as a starting point would yield ten events - a dangerously low number for modeling and prediction.

Thus, the temporal starting point for this study is March 24, 1998, for multiple reasons. First, as mentioned earlier, the federal assault weapons ban began in 1994, so examining events after its implementation offers a more contemporary examination of mass shootings. Second, using the largest modern interval creates a clean breakpoint in the event data. Since 1998, Black Swan Shootings have occurred at a historically fast pace. Further, choosing the late 1990s also correlates with the end of Mass Murder Wave 2, according to Duwe (2007). Thus, there appears to be somewhat of an intersection of policy, activity, and academic research. Third, if there is any merit to contagion and copycat behaviors associated with such extreme attacks, the Columbine High School shooting on April 20, 1999 serves as an significant milestone among Black Swan Shootings, so it should be incorporated near the beginning of the dataset (using Jonesboro as the first event makes Columbine event #3). As previously discussed, Larkin (2009) describes the legacy Columbine has created, including: the paradigm for planning and executing such an attack; inspiring future school shooters; creating a baseline casualty count for reference and comparison; and the idolized status of the attackers among vulnerable subcultures. It is in many ways the Patient Zero of Black Swan Shootings. Finally, it is worth noting that the Internet and access to worldwide content experienced a great boom in the late 1990s; access to more timely and accurate information became more realistic. Thus, this study classifies Black Swan shootings as events with either 8 killed, 13 wounded, or 15 total casualties (excluding the perpetrator(s)), occurring since 1998. There are 44 events that match these criteria.

Analytic layers

In this study, a 'layer' refers to a collection of related social factors that inform on the spatial relationships for Black Swan Shootings. They are specifically called layers, from a mapping perspective, because it is their combined overlay that is expected to strengthen any predictive power and identify communities at high-risk for Black Swan activity. There are five analytic layers: attack contagion, violence, socioeconomic, mental health, and firearm preferences. There are 13 independent variables to examine social context and location-based risk factors to potentially predict Black Swan Shootings. The layers and factors are specifically grouped this way to isolate the most relevant predictors in each grouping. This process should avoid overfitting and overweighting too many similar variables. The layers are a theme, and the social factors are the components of that theme.

The first layer of variables is related to attack contagion. This layer measures spatial clustering and temporal sprees. These are event-centric variables to determine any copycat or other similar relationships of a Black Swan Shooting on subsequent attacks over space and time. Spatially, kernel density analysis and K-means clustering is used to identify hotspots and hunting grounds of predatory behavior are applied to determine time-weighted clusters of attacks (Anselin et al., 2000; Caplan and Kennedy, 2016; Boba Santos, 2013; Paulsen et al., 2010; Hart and Lersch, 2015; Mohler et al., 2011; Chainey et al., 2008). Similar to the criminal environmental backcloth described by Brantingham and Brantingham (1993), such analyses create weighted rasters across a landscape representing desirable and undesirable crime locations (Groff and La Vigne, 2001; Paulsen and Robinson, 2009). These clusters will be used to measure how Black Swan Shootings spread, similar to the contagion modeling described by Cantor et al. (1999). Temporally, Black Swan Shootings will be organized into cycles and sprees, to identify the pace of attack occurrence. Cycles, or regularly repeated series of temporal sprees, may or may not operate on traditional calendar timelines, and can inform on patterns similar to those described by Jetter and Walker (2018). Cyclical tempo is classified as clustered (sprees), uniform (linear), and random (chaotic) patterns (Paulsen et al., 2010; Boba Santos, 2013). Sprees are short bursts of elevated activity, which can potentially correlate to the copycat-like behaviors described earlier by Carcach et al. (2002) and Towers et al. (2015). Thus, analyzing the duration, interval, and event counts of cycles and sprees can determine natural-organizing temporal activity (Lewis et al., 2011; Porter and White, 2012). This is the only analytic layer not compiled at the county and year level; the spatial and temporal calculations for contagion measures are calculated at the attack location and date levels.

The second analytic layer, violence, is measured across two variables: FBI UCR violent crime rates, and law enforcement firearm fatalities. FBI UCR violent crime includes four offense types: criminal homicide, forcible rape, robbery, and aggravated assault. These data are available annually since the 1980s nationwide from participating police jurisdictions, which include states, metropolitan statistical areas, counties, cities, colleges and universities, and tribal agencies. In this study, all rates are calculated as crimes per 100,000 residents to establish normalized crime comparisons over space and time (FBI UCR, 2018). Violent crime is a location correlate for Black Swan Shootings, as Duwe (2007), Rocque (2012), and Twemlow et al. (2002) have previously found significant relationships between crime rates and mass murder. Law enforcement firearm fatalities will also be collated from FBI UCR, using the annual Law Enforcement Killed and Assaulted (LEOKA) data for the number of officers feloniously killed and injured. This measure would ideally be coupled with citizens killed by law enforcement, but reliable open source data goes back to 2013. Violence against law enforcement is a layer in this study stemming from the previous described research from Swedler et al. (2015) and Kivisto et al. (2017). This study uses violence against law enforcement as a proxy for understanding community aggression and tension.

The third analytic layer consists of socioeconomic variables, which includes population density, racial diversity, sex, poverty, and unemployment activity. Population density is measured as residents per square mile of a county. Racial diversity is defined as the percentage of non-Hispanic Caucasian population. Sex is the percentage of the female population. Poverty is defined as the rate of people living below the U.S. Census Bureau's established threshold for minimum living expenses in a given year. Unemployment activity is measured as the percentage of a population of known unemployed adults. Unemployment data is derived from the Bureau of Labor Statistics. Socioeconomic variables are included as factors to compare to Black Swan Shootings based on the work from Duwe (2007) on economic and social correlates with mass murder, Fox and Levin (1994) on mass murder precipitating events and social connectivity, as well as Kwon and Cabrera (2017) on residential instability and mass shootings, among others. All five socioeconomic variables are measured annually by county.

The fourth analytic layer is mental health. As described by Meloy et al. (2004), Fox and Levin (1994), Lankford (2015b), and Lester (2010), mental health is a potentially significant factor to consider for mass murder. Prosocial factors in a community contribute to the mental wellness of an individual, so measuring location-centric variables such as substance abuse and quality of life to Black Swan Shootings has merit. Substance abuse is defined as drug and alcohol overdose fatalities. Quality of life is measured by suicide, which is defined as the voluntary and intentional act of taking one's own life. Both variables are derived from the National Centers for Disease Control and Prevention (CDC) and measured annually at the county levels.

The fifth analytic layer is firearm preferences. Variables for measuring firearm preference are firearm availability and political sentiment. Firearm availability is derived from firearm ownership, and firearm ownership is estimated via a proxy measure - the percentage of suicides committed with a firearm. Given the lack of clear and consistent firearm ownership records across varying state and local laws and policies, several studies have previously used such a proxy (Kleck, 2004; Azrael et al., 2004; Siegel et al., 2013; Siegel et al., 2014). As described in the literature review, high levels of ownership consistently correlate with high levels of violence; this study seeks to see if that relationship is similar with Black Swan Shootings. Data is obtained from the Multiple

Cause of Death files available from the CDC National Center for Health Statistics (Division of Vital Statistics).

Political sentiment is derived from local firearm legislation. Political sentiment in this context is an extension of elected officials supporting firearm laws that are representative of the attitudes and preferences of their constituents. Studies such as Donohue et al. (2018) and Guis (2015) demonstrate the impact firearm legislation can have on mass shootings and general violent crime, and this study builds on that research. Many methods could be used to derive public opinion (Wellford et al., 2004); laws are seemingly the most uniform over space and time. Also, as multiple studies highlighted in the literature review demonstrate, firearm laws have significant effects on the ability to acquire weapons. Using methods similar to Fleegler et al. (2013) in the analysis of state legislative scoring and firearm fatalities, in this study data for laws come from three sources of state scoring/report card measures: the Giffords Law Center; State Firearm Laws, an open source compilation from Boston University School of Public Health; and Everytown research center. Firearm legislation scorecards are compiled by state and are based on laws concerning firearm trafficking, background checks, child safety, assault weapons, and open carry in public places. The Giffords Law Center and its predecessor, the Brady Scorecard, have a comprehensive methodology for calculating their grades (Giffords Law Center, 2010). Both variables in this analytic layer are measured annually; firearm availability is at county level, and political sentiment at the state level. Table 4 provides a summary of all five analytic layers, their associated variables, the spatial and temporal precision, and the data sources.

Table 4 Panel time series analytic layers and social factor variables

Analytic Layer	Variable	Type	Variable Description	Precision		Data Sources
				Spatial	Temporal	
Black Swan Shootings	*Mass Shootings*	Event	Black Swan Shootings	Attack and Residence Counties	Annual	Black Swan Shootings dataset
Attack Contagion	*Clustering*	Event	Prior Black Swan Shooting patterns	Attack location	Attack date	Black Swan Shootings dataset
	Sprees					
Violence	*Crime*	Event	Violent Crime	County	Annual	FBI UCR
	Law enforcement violence	Event	Law Enforcement Officer Killed and Assaulted	County	Annual	FBI UCR LEOKA
Socioeconomics	*Population density*	Attribute	People per square mile	County	Annual	US Census Bureau
	Racial diversity		Non-Hispanic Caucasian population			
	Sex		Percentage of female residents			
	Poverty		People living below government baseline			
	Unemployment		Percentage of known unemployed adults			Bureau of Labor Statistics
Mental Health	*Substance abuse*	Event	Drug and alcohol overdoses fatalities	County	Annual	CDC
	Quality of life	Event	All suicides			
Firearm Preferences	*Firearm availability*	Attribute	Percentage of suicides committed by firearm	County	Annual	CDC
	Firearm political sentiment	Attribute	Firearm legislation	County (extrapolated from State)	Annual	Giffords Law Center
						Everytown
						State Firearm Law Database

Dataset

This study is based on the analysis of multiple disparate location-based data sets to detect previously unknown, non-obvious relationships in space and time to extreme mass shootings. A panel time series database is built at the county and year levels. This construct allows for measuring across all variables, including the Black Swan Shootings, for each county and year where data is available. Most spatial criminology literature recommends conducting analysis at the census tract precision level (Andresen and Malleson, 2013). However, such studies are conducted for a city, county, police jurisdiction, or other magnitude smaller than one state. For a nationwide study, county-level resolution should offer a healthy balance of more precision than a state but exponentially fewer calculations than a census tract, as counties are typically composed of multiple census tracts (Census, 2012). Table 5 highlights the variables captured for each Black Swan Shooting event that are incorporated into the panel time series database.

Table 5 Black Swan Shooting event variables

Variable	Description
Unique ID	Unique identifier for each event
Total victims	Count
Total victims anomaly	Yes/No flag for anomaly detected
Killed victims	Count
Killed victims anomaly	Yes/No flag for anomaly detected
Wounded victims	Count
Wounded victims anomaly	Yes/No flag for anomaly detected
Attack latitude	Decimal degrees coordinates
Attack longitude	Decimal degrees coordinates
Attack cluster	Analytically derived hotspot identification
Attack city	City name
Attack county	County name
Attack state	State name
Attack site count	Values ranging 1-6
Location name(s)	Attack site place name
Location type	Bar, Concert, Hotel, Manufacturing, Military Base, Movie Theatre, Office, Religious, Residence, Retail, Retirement Home, School, Street
Location area	Indoors, Outdoors
Location atmosphere	Private, Public
Location environment	Residential, Non-residential, Mixed
Date	Attack date
Year	Attack year
Sequence number	Chronological ordered number
Interval	Days since previous attack
Percent change	Percentage change between sequential intervals
Spree	Analytically derived short-term temporal clustering identification
Spree events	Number of events within identified spree
Spree duration	Number of days for identified spree
Spree Interval	Days between current and previous sprees
Cycle	Analytically derived long-term temporal clustering identification
Cycle events	Number of events within identified cycle
Cycle duration	Number of days for identified cycle
Cycle interval	Days between current and previous cycles
Firearm count	Values ranging 1-23
Weapons description	Makes and models of all weapons used
Assault weapon used	Yes/No flag for assault weapons
Offender county	County of residence at time of attack
Offender state	State of residence at time of attack
Offender race	Asian, Black, Hispanic, Middle Eastern, Native American, White
Offender gender	Male, Female
Offender age	Values ranging 15-64
Offender count	Values ranging 1-2
Prior mental health warnings	Yes/No/Unknown flag if detected
Victim-Offender relationship	Known, Unknown
Data sources	Citations for event details

These data include fields for the county of the attack location, and the county of the offender residence at the time of the attack. These are captured separately to independently analyze location-based patterns between the social factors for attack occurrence and offender residence. Prior studies have demonstrated the importance of analyzing attack locations versus residential and planning locations. Smith (2008) analyzed 60 U.S. terrorist events over 25 years, and found 44% of the perpetrators involved lived within 30 miles of their target, and among "single-issue terrorists" (which Black Swan Shooters could often be characterized), 71% of preparatory acts occurred within 12 miles of the target location. Smith et al. (2012) found residential locations for terrorists are predictors for both pre-event planning locations and eventual terrorist event locations. Smith et al. (2012) found 46% of leading activities occur within 30 miles of a terrorist's residence. Further, 57% of ancillary activities, 40% of preparatory activities, and 35% of events occur within 30 miles of the residence (Smith et al., 2012). However, given the dynamic size of counties throughout the U.S., 30 miles could be in the same or different counties. Thus, Roberts et al. (2013) examined residence and pre-attack activity locations for 20 years of domestic terrorist events and found 61% of residences and 51% of pre-event activities occur in counties different than the intended terrorism event. Therefore, this study examines both offender residence and attack location for Black Swan Shootings to see if such differences exist among these events.

Black Swan shootings and the 13 independent variables divide into two subsets: events and attributes. Events are actions that have taken place in an area and are designed to capture activities occurring in a community. There are seven event variables.

Attributes are defined as characteristics describing an area. Attribute-based variables are designed to describe the current environment at the time of the attack. There are seven attribute variables. To eliminate spatial mismatching, the variables for state-level data will be appended to each of a state's corresponding counties. There are clear benefits to using county-level data instead of states in this study. As Ayres and Donohue (2003) note, there are 3,000 counties compared to 50 states; using counties improves precision. Table 4 provides details for each variable.

Analysis plan

The data analysis plan for this study consists of seven steps. All analytic tests, measures, and techniques will be conducted using Google Sheets, Microsoft Excel, and the R Project for Statistical Computing (R Core Team, 2020). First, the Black Swan Shooting event data set will be analyzed for patterns across behavior, space, and time. Using common quantitative crime analysis techniques (Boba Santos, 2013; Paulsen et al., 2010; Hart and Lersch, 2015; Mohler et al., 2011; Chainey et al., 2008; Groff and La Vigne, 2001; Paulsen and Robinson, 2009; Cantor et al., 1999), this step will measure attack contagion and copycat activity. Behaviorally, there will be a brief analysis of casualty counts, offender characteristics, and micro-geographic location patterns, to identify any commonalities across attacks.

Spatially, this involves hotspot detection using kernel density estimation (KDE) and K-Mean clustering methods. KDE is a clustering algorithm that statistically transforms point data into clusters of spatially similar events. Based on a manually defined search radius, KDE searches across a partitioned/hashed space and identifies the number of points within range of each cell (D'Anna, 2016; Caplan and Kennedy, 2016; Boba Santos, 2013; Paulsen et al., 2010; Duong, 2007). Similar to pixilation on a television or printer, the result is a visually smoothed graph or map that identifies areas of elevated activity. K-Mean clustering is a method for partitioning data into a user-defined number of groups. After the *k* is identified, the model iterates over the entire dataset until every point is assigned to a cluster (Wagstaff et al., 2001). The application and results of these models will be discussed further in the next chapter.

Temporally, contagion and copycat activity is measured with cycle and spree calculations. Cycles and sprees are analyzed using the intervals between events. The intervals are statistically analyzed and clustered to identify natural, self-occurring significant breaks over time (Paulsen et al., 2010). Short bursts of accelerated events will be labeled as sprees, and multiple closely occurring sprees will become cycles. Thus, every event will be proscribed a cluster, spree and cycle. These measures will be analyzed for creation rates, duration, and event count to determine patterns of occurrence over space and time.

Second, the panel time series dataset is created. This dataset will be organized at the county and year levels and will range from 1998 to 2018. Fields will include an indicator for Black Swan Shooting attack, Black Swan Shooting residence, and the aforementioned social factors. For each social factor, a blend of metrics is calculated and compiled. There are four types of metrics: count; rate; percentage change; and deviation from the nationwide mean. Count refers to the raw number total for that county and year.

Rate is the count normalized by the population, expressed here as the count divided by the population, multiplied by either 100,000 or 10,000, depending on the magnitude of the variable. The percentage change is the normalized annual comparison for a county, expressed as the change for either the count or the rate for the previous year. The deviation from the nationwide mean is calculated by subtracting the nationwide count or rate mean for a variable from a given county's count or rate. A positive number represents a higher number than average, and a negative number is lower than the national average. This metric is slightly redundant, as it is expected to align closely with the rate metric. However, where this will be particularly useful is for identifying the direction and magnitude of the rate metric. Thus, the comparison to the nationwide mean metric will be useful when the rate calculation shows a significant relationship. Such a combination of calculations accounts for annual changes within a community relative to the rest of the country. Examining count or rate statically would fail to incorporate any changes over time. Table 6 identifies the calculations run for each variable.

Table 6 Social factor metric calculations compilations

Analytic Layer	Variable	Variable Description	Calculations			
			Count	Rate	Percentage change	Comparison to nationwide mean
Black Swan Shootings	*Mass Shootings*	Black Swan Shootings	Events; Casualty (Total, Killed, and Wounded)	No	No	No
Violence	*Crime*	Violent Crime	No	Per 100,000 residents	Based on Rate	Yes
	Law enforcement violence	Law Enforcement Officer Killed and Assaulted	Yes	Per 100,000 residents	No	No
Socioeconomics	*Population density*	People per square mile	No	Per square mile	Based on Rate	Yes
	Racial diversity	Non-Hispanic Caucasian population		Percentage not rate	Based on Percentage	Yes
	Sex	Percentage of female residents		Percentage not rate	Based on Percentage	Yes
	Poverty	People living below government baseline		Per 100 residents	Based on Rate	Yes
	Unemployment	Percentage of known unemployed adults		Per 100,000 residents	Based on Rate	Yes
Mental Health	*Substance abuse*	Drug and alcohol overdoses fatalities	No	Per 100,000 residents	Based on Rate	Yes
	Quality of life	All suicides				
Firearm Preferences	*Firearm availability*	Percentage of suicides committed by firearm	No	Per 100,000 residents	Based on Rate	Yes
	Firearm political sentiment	Firearm legislation	Yes	Per 100,000 residents	Based on Count	Yes

Third, a series of correlations matrices will be generated. A matrix will be created for each of the four calculations for the 11 social factors across violence, socioeconomics, mental health, and firearm preferences in the dataset. These matrices will determine the pairwise relationships for count, rate, percentage change, and comparison to the nationwide mean for the variables, to identify any unique, spurious, heteroscedastic, and collinear relationships. This will help identify redundancies and patterns independent of the Black Swan Shootings.

Fourth, three statistical tests for measuring significant differences between counties with and without Black Swan Shootings will be conducted: t test, effect sizes, and Mann-Whitney U. An independent sample, two-tailed t test will be calculated for each social analysis factor to determine if the mean of the count, the rate, the percent change from the previous year, and the comparison to the nationwide mean for counties with Black Swan Shootings is significantly different than the mean for counties without attacks. This analysis will be conducted by compiling two subsets for each analytic layer across all the years. This will help build a robust dataset rich enough to conduct statistical calculations.

For each layer, if the result of the comparison of means is determined to be statistically significant, then that layer will be deemed potentially suitable for incorporating into the predictive surface. Next, standardized effect sizes will be calculated across the same metrics (count, rate, percent change, and comparison to nationwide mean). These will be calculated for each social factor layer using Cohen’s D.

Braga and Weisburd (2012) advocate using an effect size statistic to identify the strength and direction of observed relationships. This is particularly valuable in studies with low-event counts. As Telep et al. (2014) and Coe (2002) found, sample size negatively affects variability; small studies can have higher variability in statistical power, as precision increases. Thus, measuring the effect sizes of each factor in this study will help determine the magnitude of the difference between counties with Black Swan Shootings and counties without such attacks for each year and the overall dataset. The analytic layer(s) found to have the greatest effect, described here as the highest percentiles using a natural break in the results, for counties with Black Swan shootings will be deemed potentially suitable for incorporating into the predictive surface. Finally, a Mann-Whitney U test will be conducted. Such a measure is well suited for this study, as it is designed to measure the medians of nonparametric data. This test will be used to assess the similarity of distributions for counties with and without Black Swan Shootings, across each metric for all social factors. Mann-Whitney U will help determine the stochasticity of each variable (Ford, 2017). The factors exhibiting the greatest difference in distribution and statistical significance will be deemed potential variables for the predictive model. Since the event counts are so low, this combination of three tests should provide enough statistical rigor to identify valid relationships and explore all facets of every variable as much as possible. As King and Zang (2009) found, it is not necessary to work with sample populations when the total population is so small. Rather, it is necessary to thoroughly and robustly analyze and enrich those events. This set of tests seeks to accomplish that goal.

Fifth will be the selection of analytic layers to incorporate in the predictive model. This step will use the significant results from each of the calculations in the previous step. This process of leveraging the appropriate statistical measures across each of the different metrics is designed to account for the rarity of Black Swan Shootings. As King and Zeng (2001) described, when working with small data samples the process of enriching attributes increases the depth of the events. Thus, social factor variable selection is a two-step process. First, initial selection is conducted as follows:

- Variables with a statistically significant t test at the $p < 0.05$ level are included. This p-value threshold is the general standard in social science research.
- Variables with a Cohen's D effect size of 0.30 or greater are included. Given the exceptionally large differences in sample sizes for counties with Black Swan activity versus those without, the effect size is measured with an assumption of unequal variances. Cohen's D scores can be tiered as small, medium, and large. Small effects are approximately 0.2; medium effects are near 0.5; and large effects are greater than 0.8. Thus, this threshold is chosen as it is deemed just above the "low score" established in prior studies (Braga and Weisburd, 2012; Coe, 2002).
- Variables with a statistically significant Mann-Whitney U test at the $p < 0.05$ level are included. This p-value is considered a universally accepted standard as well (Ford, 2017).

Thus, if a social factor variable was determined to not have a statistically significant difference in means, nor have a significantly high effect size, nor have a significantly different distribution, it is excluded from further analysis. This process of filtering analytic layers is key; even if the predictive model does not perform well, this study will have identified which social factors have a distinctly unique relationship with the locations of Black Swan Shooting attacks and residences. Second, this set of included variables is further analyzed. Each of the remaining variables are analyzed for strength of statistical significance, diversity, and relevance to Black Swan Shooting behaviors. Variables are compared for the number of statistical tests that were significant, as well as how similar they are from the correlation tests in step three of this plan. Similar variables are de-duplicated and prioritized by the variables hitting on the higher number of statistical tests. This step results in a set of variables to model further.

Sixth, logistic regression models are built. Logistic regression is the preferred model type for this study because the outcome attempting to be predicted, whether or not a Black Swan Shooting occurred, is dichotomous. The panel time series dataset, as constructed, has a binomial dependent variable for Black Swan attack counties and Black Swan residence counties. The predictor variables are determined from the results of step five. Thus, the predicted values for each of those variables are the probability of a Black Swan Shooting occurring (for attack and residence) in a given county.

Seventh, the model's performance will be tested for each year. This test will measure the accuracy and precision of the predictions by the count and locations of attack and residence. This will result in a yearly performance table used to assess the utility of the regression model. The specifics of this metric will be described in the next section. This will allow for a rigorous assessment of the value of such techniques and the ability to model Black Swan Shootings. If these tests are deemed successful, future studies will potentially be able to apply this model to future years of attacks (and in particular when the analytic layer data becomes available).

Testing and metrics

As described in the introduction, most criminal justice research involving prediction is centered around offenders, sentencing, and recidivism. It is not typically event centric. An event-based criminology prediction should incorporate accuracy and precision over space and time, and the evaluation of such predictions should measure the same components. Kahneman and Tversky (1973) identified three types of information relevant to prior probabilities and known evidence of statistical predictions: background information; specific event details; and the expected accuracy of the prediction. Accuracy is the degree to which something is *correct* (Chaiken et al., 1994). Accuracy should be a calculation of the window of a future event occurring in space and time. The degree of accuracy is a function of the amount of background and event information available. Precision is the size of the spatial and temporal area predicted. In the law enforcement and national security sectors, prediction precision is paramount for maximum utility. A vague prediction increases its accuracy but sacrifices the ability to be operationalized. Thus, tactically relevant predictions should have a precision in the 95th percentile or greater of the overall space and time of a series (Ardohain, 2016).

Predictions can be relatively easy to make but assessing the performance of predictions is rarely documented and shared in this community. The paucity of research surrounding law enforcement and national security prediction assessment lies in a lack of consistent and uniform measurement. A purely binary measure - indicating a simple yes/no for future event occurrence - seems to be the least complicated; however, it eliminates much of the nuance and potential degrees of correctness. Within the realm of the United States Department of Defense, analytic methodologies were developed to identify, analyze, and resolve tactical patterns of events. The Joint Improvised Explosive Device Defeat Organization (JIEDDO) leveraged the theory and methods of law enforcement crime analysis to better address the decision-making processes of insurgent attackers (Toffler Associates, 2011). As such, the COIC Pattern Analysis Team (CPAT) within JIEDDO developed predictive models leveraging mathematics, psychology, and other related sciences for global improvised explosive device (IED) activity (Shankar, 2008; Ardohain, 2016). The results of these analyses were quantitatively measured for predictive performance. The JIEDDO CPAT team served as one of the few known examples of a group using crime analysis methodologies actively conducting predictions on serial, predatory behaviors of low-count events, and routinely measuring their performance. Analysis of those predictions suggest using the predicted times and locations furthest from the actual next event, to systematically over-penalize the performance. By designing a metric that intentionally scrutinizes the analysis, operators (whether they are law enforcement, military, or others) are more likely to accept the results as valid. Using the lessons learned from measuring such predictions, the metric shown in Equation 1 is proposed for assessing the precision and accuracy of Black Swan Shooting annual prediction counts and locations.

Equation 1 Black Swan prediction metric

$$\frac{1}{3}\left(Precision_{Space}\right) + \frac{1}{3}\left(Accuracy_{Space}\right) + \frac{1}{3}\left(Accuracy_{Time}\right)$$

$$Precision_{Space} = \left(1 - \left(\frac{County_P}{County_C}\right)\right)$$

$$Accuracy_{Space} = \left(\left(\frac{\min(\sum Attack_D \mid County_C)}{\max(\sum Attack_D \mid County_C)}\right)\right)$$

$$Accuracy_{Time} = \begin{cases} 1 & if\ Attack_P = Attack_A \\ \dfrac{min(Attack_P \mid Attack_A)}{max(Attack_P \mid Attack_A)} & otherwise \end{cases}$$

$\boldsymbol{Attack_P}$ = *count of predicted Black Swan Shootings for a given time period*

$\boldsymbol{Attack_A}$ = *count of actual Black Swan Shootings for a given time period*

$\boldsymbol{County_C}$ = *count of all U.S. Counties or States*

$\boldsymbol{County_P}$ = *count of unique U.S. Counties predicted for a Black Swan Shootings for a given time period*

$\boldsymbol{Attack_D}$ = *count of states in between the predicted county and a Black Swan Shooting for a given time period*

Results in a score between 0 and 1

A few points of discussion on this metric. First, this formula is designed to account for accuracy (correctness) and precision (exactness) over space and time. Spatial accuracy is measured as the number of counties separating a high-risk county from a county experiencing an attack, calculated as the Euclidean (straight-line) distance for each high-risk county to its nearest actual attack. Distance measures are intentionally reduced to a count of counties to normalize for county sizes. Spatial precision is measured as the number of counties at highest risk in a given year, compared to the total number of U.S. counties. Temporal accuracy is a comparison of the count of highest risk counties to the count of counties experiencing attacks, in a given year. Temporal precision is controlled by using the highest risk counties each year. While there are four components to analyzing model performance (spatial accuracy, spatial precision, temporal accuracy, temporal precision), three (spatial accuracy, spatial precision, and temporal accuracy) are measured, and one (temporal precision) is inferred. Thus, each component of the calculation is weighted equally (one-third).

Second, the inputs are the count and location of counties at the highest risk for a Black Swan Shooting in a given year. Third, these high-risk counties are derived from the tiered regression model outputs for a given year. Fourth, this model intentionally penalizes strongly for zeros. This means that if no attacks occur in a given year, but the prediction is for greater than zero, the score is significantly impacted. Fifth, a table of performance for each year will be generated, where the results of the metric are a value between 0 and 1. Retro-predictions will be calculated for each year since 1999, compiled and analyzed for trends and patterns. Based on the performance over time, a natural cutoff and basic averages will be identified. Retesting will be conducted on the predictability of combinations of layers. This process will be applied to both attack and residence counties independently.

ANALYSIS

This chapter addresses the analysis and results of this study in six section. First, a discussion on the Black Swan Shooting patterns, independent of the social factor variables, to include any spatial clustering and temporal spree activity. This includes any indications of copycat and contagion effects. Second, an analysis of the social factor variables, independent of the Black Swan Shootings. This addresses any correlations across analytic layers. Third, the results of the series of statistical measures applied to the social factor variables. This process compares the means of counties with Black Swan Shooting attacks and residences, to counties without such activity, across multiple metrics. Fourth, a description of the analytic layers selected based on the results of the statistical tests. This identifies the layers with the strongest relationships to Black Swan Shootings. Fifth, the creation of the logistic regression models for Black Swan attack counties and Black Swan residence counties. Sixth, an analysis of the model's performance over time based on the aforementioned metric for measuring number and location of attacks and residences.

Black Swan event patterns

With the event data established in the methodology chapter, some basic analysis can be conducted to determine any trends, patterns, and correlations of interest.

Understanding Black Swan attack patterns can best be understood across three dimensions: behavior, space, and time.

Behavior

Analyzing the 44 Black Swan Shootings revealed patterns of casualty counts, offender demographics, and attack components. When examining casualty counts, Table 7 shows that mid-to-lower casualty count events are the most frequent. Killed counts spike at 8-9 victims, with 9 or less killed victims accounting for 61% of all Black Swan Shootings. Further, 30% of Black Swan Shootings involve 11 or more killed victims, and 11% of events have 17 or more people killed. Wounded victim counts also cluster, but differently. Thirty percent of Black Swan Shootings have two or fewer wounded victims, and 13% of events have 17 or more wounded. The spike here is at 13 wounded, of which 5 events (11%) experienced. Finally, 45% of all Black Swan Shootings (20 events) have 8-15 total victims. Among the highest total victim counts, 32% of attacks have 20 or more total victims, and each of those 15 events have unique victim counts.

Table 7 Black Swan Shooting event counts and percentages, by total, killed, and wounded victims

Killed Victims	Event Count	Percent of attacks	Cumulative Percent
2	2	5%	5%
3	2	5%	9%
4	1	2%	11%
5	2	5%	16%
6	4	9%	25%
7	1	2%	27%
8	8	18%	45%
9	7	16%	61%
10	3	7%	68%
11	1	2%	70%
12	2	5%	75%
13	3	7%	82%
14	2	5%	86%
17	1	2%	89%
27	1	2%	91%
28	1	2%	93%
32	1	2%	95%
49	1	2%	98%
59	1	2%	100%

Wounded Victims	Event Count	Percent of attacks	Cumulative Percent
0	4	9%	9%
1	4	9%	18%
2	5	11%	30%
3	1	2%	32%
4	2	5%	36%
5	1	2%	39%
7	2	5%	43%
8	2	5%	48%
9	1	2%	50%
10	1	2%	52%
11	1	2%	55%
12	2	5%	59%
13	5	11%	70%
14	2	5%	75%
17	1	2%	77%
20	1	2%	80%
21	2	5%	84%
23	1	2%	86%
24	1	2%	89%
25	1	2%	91%
30	1	2%	93%
53	1	2%	95%
70	1	2%	98%
441	1	2%	100%

Total Victims	Event Count	Percent of attacks	Cumulative Percent
8	3	7%	7%
9	2	5%	11%
10	4	9%	20%
11	3	7%	27%
13	1	2%	30%
15	7	16%	45%
16	2	5%	50%
17	1	2%	52%
18	4	9%	61%
19	2	5%	66%
20	1	2%	68%
22	1	2%	70%
23	1	2%	73%
27	1	2%	75%
29	1	2%	77%
30	1	2%	80%
34	1	2%	82%
35	1	2%	84%
37	1	2%	86%
43	1	2%	89%
47	1	2%	91%
55	1	2%	93%
82	1	2%	95%
102	1	2%	98%
500	1	2%	100%

Related to victim counts, multiple studies reviewed discuss a prior victim-offender relationship. Among Black Swan Shootings, 29 of the 44 attacks (66%) involve a perpetrator that had an existing relationship with the victims. This relationship is most often a co-worker or classmate.

Regarding demographics, offender patterns largely follow prior research on perpetrators of mass violence. There are 47 total perpetrators across the 44 Black Swan Shootings. The average number of offenders per attack is 1; 41 attacks are a single perpetrator, and three attacks have two. Of the 47 attackers, 45 are males. A female is the sole perpetrator in one attack (1/30/2006 in Goleta, California), and a partner in another (12/2/2015 in San Bernardino, California). Examining the attacks by race, 32 attackers are Caucasian (68%), 7 are African American (15%), 4 are Middle Eastern (9%), 2 are Asian (4%), and 1 each for Hispanic, and Native American. This runs slightly contrary to the Fox and Levin (1998) assessment that mass murderers fall along generic population distributions. These are also slightly different percentages than Dillon (2013) found; in particular there are higher numbers of black offenders among Black Swan Shooters. The average age is 30 years old, with a minimum of 11 and maximum of 64. The standard deviation among the 47 ages is 12; thus, 95% of the perpetrators are between 18 and 42 years old. This average is slightly younger than the mid-30s age identified by most studies on mass violence and influenced by the number of school shootings involving juveniles. Nineteen percent of offenders are juveniles, 30% are in their 20s, 19% are in their 30s, and 25% are older than 40.

Regarding the dynamics of the attack, patterns exist among the number of sites and location types. The average number of target sites per attack 1.4, with a high of six (Geneva, Alabama, 3/10/2009). Of the 44 Black Swan Shootings, 34 are single-location attacks, four attacks occurred at two sites, and five attacks occurred across three sites.

Attack locations were coded across several fields for each site involved in an event. The fields captured include site environment, atmosphere, area, and type. Environment is coded as residential, non-residential, and mixed. Of the 44 Black Swan Shootings, 34 attacks (77%) occur in non-residential settings; four attacks occur in an exclusively residential environment, and six are mixed locations. Similarly, the atmosphere field is coded as public, private, and mixed. Public atmosphere is considered any venue the general population can have equal opportunity to access, without special permissions. Fourteen attacks (32%) occur in such areas, which include shopping malls and retail locations, bars and nightclubs, movie theatres, concerts, religious institutions, and open air/street locations. Private atmospheres are considered to be locations requiring special access or permissions, where the rights of the general population are limited or restricted. These include schools, residences, non-retail commercial locations such as office and manufacturing facilities, and military bases. Labeling schools as a private location is a complicated decision; however, ultimately it was determined that access to the facility, even for a public school, is not completely open and available to the entire population. Twenty-five attacks (57%) occur in private spaces. Mixed atmosphere describes multi-site attacks occurring in public and private atmospheres; 5 attacks (11%) fit these criteria. Location area is coded as indoors, outdoors, and mixed. Of the 44 Black Swan Shootings, 36 attacks (82%) occur in exclusively indoor settings, three occur in outdoor-only venues, and five attacks are multi-site events occurring in indoor and outdoor areas. Finally, location type is the description of the venue. Of the 44 attacks, there were 54 unique venues which fall into 14 categories. Among location types, Black Swan Shootings occur most often at schools (14 attack sites, 26%), residences (10 attack sites, 19%), followed by manufacturing plants and open-air/street locations (5 attack sites each), and then religious institutions (four attack sites). Table 8 provides counts for all the unique combinations of environment, atmosphere, area, and location type.

Table 8 Black Swan Shootings categorized by location details

Location Type	Count	Environment	Atmosphere	Area
Bar	2	Non-residential	Public	Indoors
Concert	1	Non-residential	Public	Outdoors
Hotel	1	Non-residential	Public	Indoors
Manufacturing	5	Non-residential	Private	Indoors
Military base	3	Non-residential	Private	Indoors
Movie Theatre	1	Non-residential	Public	Indoors
Office	3	Non-residential	Public	Indoors
Religious	4	Non-residential	Public	Indoors
Residence	10	Residential	Private	Indoors
Retail	3	Non-residential	Public	Indoors
Retirement Home	1	Non-residential	Private	Indoors
School	14	Non-residential	Private	Indoors
Street	6	Non-residential	Public	Outdoors
Sum	*54*	*81% Non-residential, 19% Residential*	*61% Private, 39% Public*	*87% Indoors, 13% Outdoors*

Space

Aggregated spatial analysis is conducted by residence and by attack locations at the county and state levels. Offenders resided in 41 unique counties within 25 states at the time of their attacks. Three of those 41 counties (Arapahoe, Colorado; Bell, Texas; Tarrant, Texas) had two offenders; the other 38 had one. Eight of the 25 states had more than one offender, with Texas (8) and California (7) having the most. The other six states with greater than one offender (Colorado, Connecticut, Florida, Mississippi, Oregon, and Virginia) all had two. Attack locations occurred in 41 unique counties across 24 states and the District of Columbia. Of the 44 Black Swan Shootings, 34 (77%) had the same residence and attack county. Three of the 41 counties had repeat attacks. By state, eight states had more than one attack, with California (8) and Texas (7) leading the way, followed by Colorado, Connecticut, Florida, Mississippi, Oregon, and Virginia. Table 9 highlights these trends for by county and state and residence and attack locations.

Table 9 Black Swan Shooting offender residence and attack location by county and state

State	Count		County	Count	
	Residence	Attack		Residence	Attack
Alabama	1	1	Coffee	1	0
			Geneva	0	1
Arizona	1	1	Pima	1	1
Arkansas	1	1	Craighead	1	1
California	7	8	Los Angeles	1	1
			Orange	1	1
			San Bernardino	1	1
			San Diego	1	1
			Santa Barbara	1	2
			Tehama	1	1
			Ventura	1	1
Colorado	2	2	Arapahoe	2	2
Connecticut	2	2	Fairfield	1	1
			Hartford	1	1
District of Colombia	0	1	Washington	0	1
Florida	2	2	Broward	1	1
			Orange	0	1
			St Lucie	1	0
Georgia	1	1	Fulton	0	1
			Henry	1	0
Illinois	1	1	Champaign	1	0
			DeKalb	0	1
Kansas	1	1	Harvey	1	1
Kentucky	1	1	Marshall	1	1
Michigan	1	1	Kent	1	1
Minnesota	1	1	Beltrami	1	1
Mississippi	2	2	Kemper	1	0
			Lauderdale	1	1
			Lincoln	0	1
Nebraska	1	1	Douglas	0	1
			Sarpy	1	0
Nevada	1	1	Clark	1	1
New Mexico	1	1	Cibola	1	0
New York	1	1	Broome	1	1
North Carolina	1	1	Moore	1	1
Oregon	2	2	Douglas	1	1
			Lane	1	1
Pennsylvania	1	1	Allegheny	1	1
South Carolina	1	1	Charleston	0	1
			Lexington	1	0
Texas	8	7	Bell	2	2
			Collin	1	1
			Comal	1	0
			Dallas	1	1
			Galveston	1	1
			Tarrant	2	1
			Wilson	0	1
Virginia	2	2	Appomattox	1	1
			Montgomery	1	1
Wisconsin	1	1	Waukesha	1	1

However, examining by county and state alone is not sufficient for determining clusters for any copycat and contagion effects. Instead, using the exact attack locations (by the address level when available, the city when not) allows for greater precision, and the ability to cut across inconsistent geometries and artificial boundaries. The process of hotspot detection involves mapping sequential events and calculating two-dimensional density plots. A cluster is when points are determined to be significantly closer than if they were randomly distributed. This method is preferred over county and state counts because of its reflexive, self-organizing nature; the nearest neighbor distances change with each new event. For this process, hotspots are calculated in two ways. First, a manual kernel density estimation (KDE) is run to create contour bands. These contour bands measure closeness, and events in the same contour band are deemed to be part of the same hotspot. Figure 10 shows iterations of KDE for 11, 22, and 44 attacks.

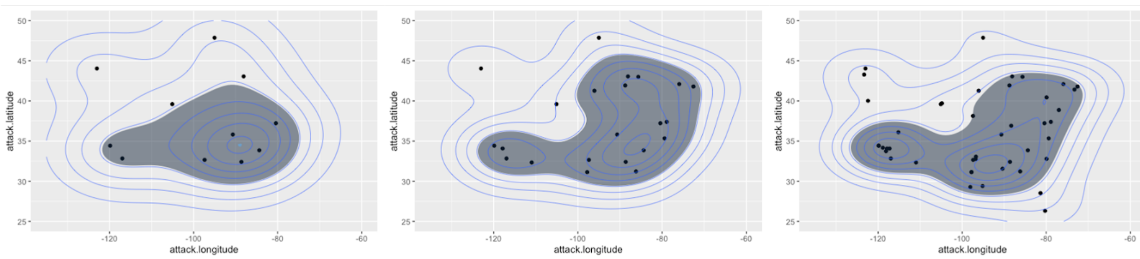


Figure 10 Black Swan Shooting kernel density estimation hotspots at 11, 22, and 44 attacks

This analysis was run starting with the second attack, to identify the number of clusters created over time. After multiple runs and evaluations, there were 23 recorded hotspots. While 23 hotspots across 44 events implies roughly two attacks per cluster, that is not the case. Rather, this process demonstrated the highly clustered nature of attacks. While several events occur in relative isolation, many attacks cluster into small geographic areas. However, this process was relatively manual, and involved a fair amount of interpolation. While this method naturally self-organizes and recalculates the data with each new event, the nature of the point distribution and shape of the country creates challenges. Contour bands have a tendency to stretch across large areas, and for a study at this scale, such analysis has the potential to create "hotspots" the size of some states. While KDE initially demonstrates that attacks cluster, a more precise calculation should better identify these hotspots.

Thus, a K-Means clustering algorithm is also leveraged. This process is more reliable, more statistically rigorous, and able to build from the baseline KDE hotspot detection. K-means calculations began with the third attack and iterated until the 44th. While K-Means does not involve counting events per contour band, there is still an arbitrary decision made for determining the number of clusters. For this study, subjectivity was removed by implementing the elbow method, originally pioneered by Thorndike (1953). This process involves graphing the within groups sum of squares values. That graph is used to determine the earliest, most usable sudden change; that value literally creates an elbow in the graph and becomes the optimal number of clusters to create. In Figure 26, the graph is the sum of squares for the first 14 attacks. The sudden change, apart from the obvious drop between 1 and 2 (which is consistent and expected), demonstrates ‘7’ is where the elbow occurs. Once the natural breakpoint is found, that is deemed the optimal number of clusters.

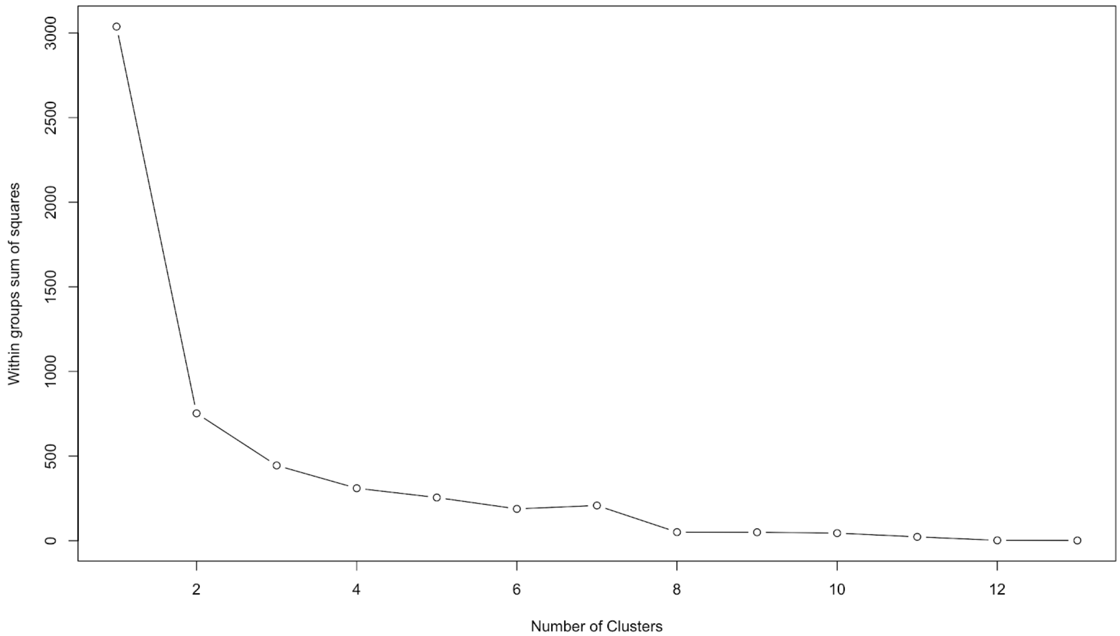


Figure 11 Black Swan Shooting attack location within groups sum of squares values, at 14 attacks

So, for each iteration, the elbow in the graph dictates how many clusters should be created. Those clusters are then calculated and assigned to the attack points. A map is produced, color coding the points by cluster. Figure 12 shows the seven hotspots identified in the first 14 attacks.

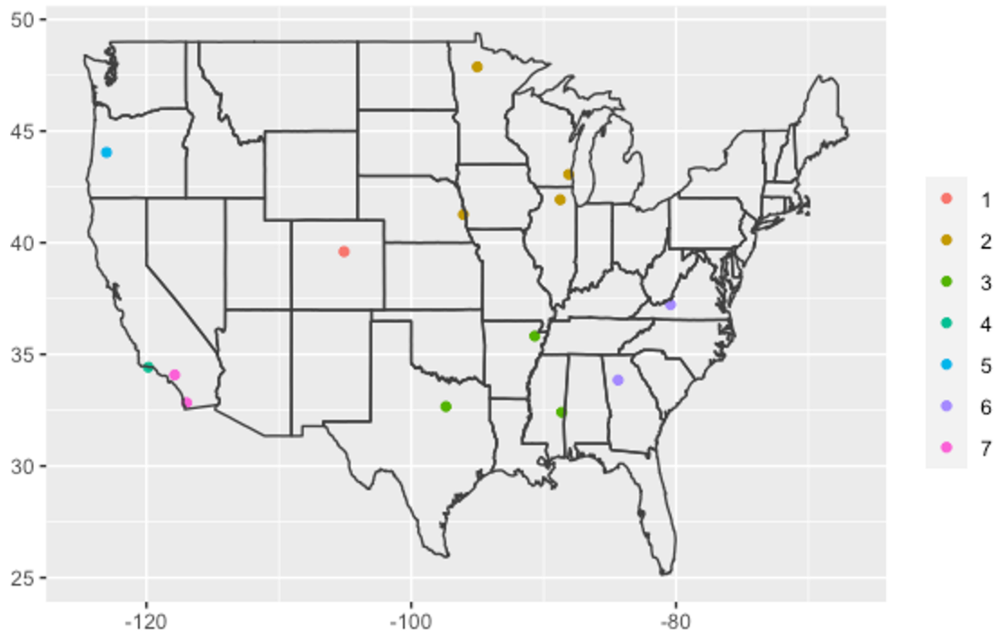


Figure 12 Black Swan Shooting attack hotspots using k-means clustering, at 14 attacks

There are two different ways to potentially determine a temporal start for this Because this is a dynamic process, with each new attack hotspots can emerge, dissolve, expand, or contract. So, the number of hotspots can actually decrease over time. For example, attack 17 shrunk the number of clusters from nine to eight. It can also lead to irregular jumps in multiple new hotspots. For example, at attack 19, activity went from nine hotspots to 12. Using the elbow method ensures rigorous testing and dynamic hotspot attribution. Table 10 highlights the changes in the number of hotspots for each new attack. Scanning for patterns of hotspot count over time, the most noteworthy stretch of relatively consistent hotspots is during attacks 19-33. During that 15-attack window, which accounts for 34% of all Black Swan Shootings, the number of hotspots stays between 11-13. And even beyond that, attacks 34-42 grow to 13-15 hotspots. Thus, during a 24-attack period (55% of all events), there are a constant 10-15 hotspots across the country.

Table 10 Black Swan Shooting attack hotspots using k-means clustering over time

Attack	Hotspot Count	Attacks Per Hotspot	New Hotspot
1	n/a	1.0	n/a
2		1.0	
3	2	1.5	N
4	3	1.3	Y
5		1.7	N
6		2.0	N
7	5	1.4	Y
8		1.6	N
9		1.8	N
10		2.0	N
11		2.2	N
12	6	2.0	Y
13		2.2	N
14	7	2.0	Y
15	9	1.7	Y
16		1.8	N
17	8	2.1	N
18	9	2.0	N
19	12	1.6	Y
20	11	1.8	N
21	10	2.1	N
22	11	2.0	N
23	12	1.9	N
24	13	1.8	Y
25	12	2.1	N
26		2.2	N
27		2.3	N
28	13	2.2	N
29	11	2.6	N
30	10	3.0	N
31	12	2.6	N
32		2.7	N
33		2.8	N
34	13	2.6	N
35	14	2.5	Y
36		2.6	N
37	13	2.8	N
38	15	2.5	Y
39	13	3.0	N
40	14	2.9	N
41		2.9	N
42	15	2.8	N
43	16	2.7	Y
44	17	2.6	Y

Overall, the 44 Black Swan Shootings cluster into 17 distinct areas, as illustrated in Figure 13. These 17 areas exist in geographically unique and distinct areas of the country. While some of these clusters are seemingly near each other and could be merged into larger geographic units (for example, the Southeastern United States), these 17 areas are distinct enough to provide insights into how events occur over time. Also, while it is tempting to examine the point distribution with the naked eye and question some of the cluster sizing and allocation, it is not recommended. Ultimately, a flat two-dimensional rendering of a landmass the size of the United States from a spheroid-shape planet results in distortion. This is true for all projections, and while steps are taken in this study to reduce edge error and produce accurate visualizations, it is not perfect. Thus, distortion aside, the points are color coded based on their cluster.

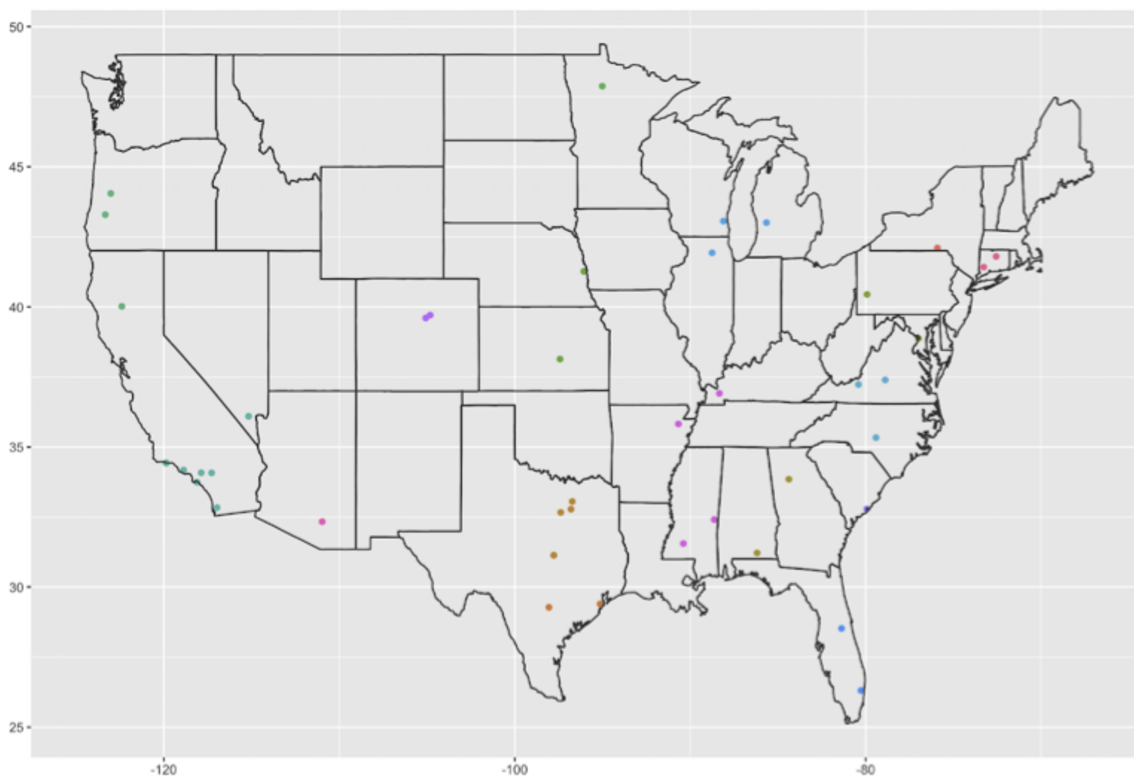


Figure 13 Black Swan Shooting hotspots using k-means clustering, at 44 attacks

Further, Figure 14 labels the 17 hotspots based on their origin. Thus, hotspot A encompasses four attacks, including the first Black Swan Shooting in the dataset: Jonesboro, Arkansas on 3/24/1998. The last hotspot, Q, was created with the Sutherland Springs shooting on 11/5/2017. Figure 14 demonstrates how counting events by states misses precise clustering. Eight of the 17 hotspots include events from multiple states, including hotspot A (four states) and hotspot G (three states). Figure 14 also depicts where clusters do *not* exist. Most notably, there are no Black Swan Shootings north of Connecticut and central New York, leaving Massachusetts, Rhode Island, Vermont, New Hampshire, and Maine without activity. There are also no Black Swan Shootings across much of the northwest of the country, starting in central Wisconsin and continuing through Iowa, most of Minnesota, North Dakota, South Dakota, most of Nebraska, Wyoming, Montana, Idaho, Utah, and Washington. Other smaller pockets of no activity also exist in the southwest and central areas of the country, but the northeast and northwest are the most noticeable.

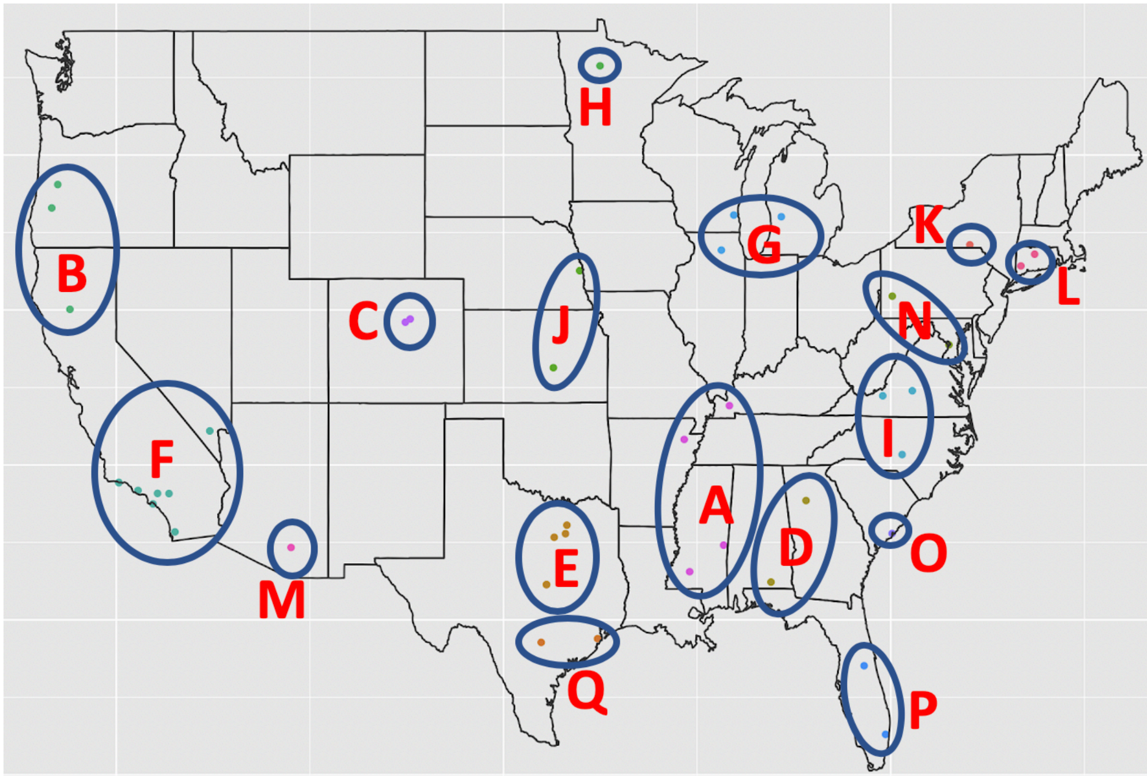


Figure 14 Black Swan Shooting hotspots, labeled A-Q based on origination

Further, by examining the origin of each hotspot for creation, duration, and return rates, clustering patterns emerge. Hotspot creation follows irregular patterns over time. On average, new hotspots are created every 2.2 events. But these are not consistent intervals. The first six attacks each created new hotspots, A-F. After those first six attacks, new hotspots are created after 1-4 new attacks; most often there are 2-4 attacks in old hotspots, followed by a new hotspot being created. Figure 15 graphs the events that create new hotspots and visualizes the slowing rate of creation over time.

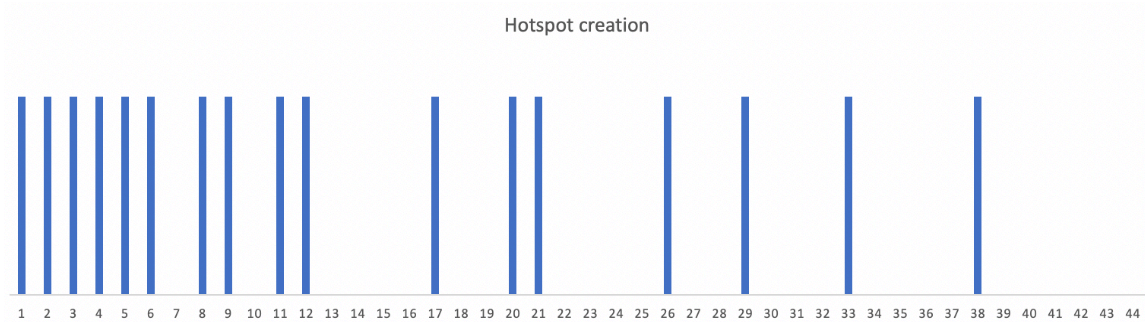


Figure 15 Black Swan Shooting hotspot creation over time

On average, each of the 17 hotspots has 2.6 attacks. The most active area is hotspot F, in the southern California/Nevada area, with eight attacks. This further highlights the difference in using hotspots versus counting by state; California has 8 attacks across two dispersed hotspots. The attack in northern California is much closer to the two attacks in Oregon, making it a part of hotspot B; rather than being grouped with the seven attacks in the Los Angeles/San Diego area. Similarly, the seven attacks in Texas cluster into hotspots E and Q. As Figure 32 indicates, there are four single-attack hotspots (23.5% of hotspots). This begs the question: are single events an actual hotspot? Yes, this study contends, for multiple reasons. First, clustering is able to identify naturally organized spatial relationships amongst the attack locations. An event that is not distinctly near other events is thus in its own cluster. A cluster is a cluster, regardless of the number of points it consists of. Second, when working with relatively small datasets, the goal is to capture as much information about the events as possible. Removing single-event clusters as hotspots would negatively impact the analysis of the dataset. Third, every hotspot begins as a single event. It is the analysis of the change of that cluster over time that matters. Thus, these four single-event hotspots are potentially the beginning of a larger future hotspot. These attacks occur in hotspot H (attack 9, on 3/21/05), hotspot K (attack 17, on 4/3/09), hotspot M (attack 21, on 1/8/11), and hotspot O (attack 29, on 6/17/15). These attacks all occur 2-4 years apart, and after the return to existing clusters had begun. This also indicates that all new attacks since Summer 2015 have occurred near a prior attack, and in an existing hotspot.

Regarding duration, on average a hotspot remains active for over seven years, across 16 attacks. Most of the early hotspots remain active for nearly the entire time range. The first six hotspots created have each occurred across 79% of the entire attack timeline, with four of the six lasting for greater than 85% of the life cycle. There is a strong correlation (0.80) between the number of events in a hotspot, and its duration. Related, return rates have decreased as new hotspots have emerged. Return rates have ranged from two events after a hotspot is created, to as much as 28 events later. Most returns to an existing hotspot occur 2-13 attacks later, and on average the return rate across all hotspots is eight events. There is a weak correlation (0.21) between the number of events in a hotspot, and the average return rate. In general, older hotspots have return rates in the 11-21 event range, and newer hotspots have return rates in the 4-8 event range. The most recent new hotspot occurred with attack 38 (hotspot Q in south Texas). Four attacks later, this area was revisited. A similar pattern happened with the previous most recently created area (hotspot P in south Florida). Hotspot P began with attack 33 and was revisited on Attack 41. New hotspots are still emerging and are not expected to have just one attack, which is another reason to count single-event clusters as hotspots. Table 11 shows the number of events in each hotspot, duration statistics, and return rates.

The oldest and newest hotspots are still active, but the hotspots created in the middle of the event timeline have seemingly gone fallow. All of the first six hotspots have lasted for nearly 10 years or more, with four of the six enduring for 18 of the 21 years of attacks. This is indicative of spatial clustering occurring many years apart. Hotspots G-M all consist of 1-3 events, began between events 8-21, ended between events 9-32, averaged a duration of eight events and 2.8 years, with an average return rate of five events. All of these calculations are notably smaller than the earliest and latest hotspots (A-F, and N-Q). Hotspots G-M appear to be short bursts of activity that have not been recently revisited.

As new Black Swan Shootings occur after this study, there are two potential outcomes for these areas. First, these hotspots may remain permanently inactive. If so, that would eliminate much of the Black Swan Shooting activity in the central/Midwest portions of the country. It would be interesting to measure any changes in those areas after Black Swan Shootings had occurred, to identify any potential disruption, deterrence, or displacement effects. Second, hotspots G-M may be inactive temporarily, and with new attacks in those areas those hotspots would start to resemble the size and duration of the earlier hotspots A-F.

Table 11 Black Swan Shooting hotspot statistics

Hotspot	State(s)	Event Count	First Event	Last Event	Event Duration	Duration (years)	Percent of total timeline	Average return rate
A	Mississippi, Arkansas, Kentucky	4	1	40	40	19.9	96%	13.0
B	Oregon, California	4	2	39	38	19.5	94%	12.3
C	Colorado	2	3	24	22	13.3	64%	21.0
D	Alabama, Georgia	2	4	15	12	9.6	47%	11.0
E	Texas	5	5	36	32	18.0	87%	7.8
F	California, Nevada	7	6	44	39	17.7	86%	6.3
G	Illinois, Michigan, Wisconsin	3	8	22	15	6.3	31%	7.0
H	Minnesota	1	9	9	1	0.0	0%	0.0
I	Virginia, North Carolina	3	11	19	9	2.8	13%	4.0
J	Nebraska, Kansas	2	12	32	21	8.2	40%	20.0
K	New York	1	17	17	1	0.0	0%	0.0
L	Connecticut	2	20	25	6	2.4	11%	5.0
M	Arizona	1	21	21	1	0.0	0%	0.0
N	Pennsylvania, Washington DC	2	26	43	18	5.1	25%	17.0
O	South Carolina	1	29	29	1	0.0	0%	0.0
P	Florida	2	33	41	9	1.7	8%	8.0
Q	Texas	2	38	42	5	0.5	3%	4.0
Average	**1.5**	**2.6**	**14**	**29**	**16**	**7.4**	**36%**	**8.0**

Given these Black Swan Shooting spatial patterns, a new attack is expected to fall within an existing hotspot rather than create a new one. Analyzing hotspot durations and return rates can identify which of the 17 hotspots remain viable for future attacks. By applying each of the prior return rates to the most recent attack in each hotspot, there are nine hotspots at the highest risk for future attacks. Table 12 identifies the nine hotspots and the potential future event number(s). Figure 16 highlights these nine hotspots on the map. In simplest terms, the map indicates the areas for highest risk are California, Colorado, Kansas/Nebraska, Texas, Mississippi/Arkansas/Tennessee/ Kentucky, Pennsylvania/Virginia/Maryland, and Florida. There is one caveat. The four single-event hotspots (hotspots H, K, M, and O) are deemed not viable for now, because they lack a return rate. By applying both the average return rate (8 events) and the average first return rate (11 events) to each of these hotspots, none of them are still active. However, given that return rates for other hotspots have been as high as 28, these areas do have potential for future activity, but are currently too unstable to include in the "highest risk" hotspots.

Table 12 Black Swan Shooting highest risk hotspots for future events

Hotspot	State(s)	Event Count	First Event	Last Event	Prior return rates	Potential next event(s)
A	Mississippi, Arkansas, Kentucky	4	1	40	6, 28, 5	45, 46, 68
B	Oregon, California	4	2	39	26, 2, 9	48, 65
C	Colorado	2	3	24	21	45
D	*Alabama, Georgia*	*2*	*4*	*15*	*11*	
E	Texas	5	5	36	13, 9, 7, 2	45, 49
F	California, Nevada	7	6	44	4, 4, 9, 8, 6, 7	48, 50-53
G	*Illinois, Michigan, Wisconsin*	*3*	*8*	*22*	*5, 9*	
H	*Minnesota*	*1*	*9*	*9*	*0*	
I	*Virginia, North Carolina*	*3*	*11*	*19*	*5, 3*	
J	Nebraska, Kansas	2	12	32	20	52
K	*New York*	*1*	*17*	*17*	*0*	
L	*Connecticut*	*2*	*20*	*25*	*5*	
M	*Arizona*	*1*	*21*	*21*	*0*	
N	Pennsylvania, Washington DC	2	26	43	17	60
O	*South Carolina*	*1*	*29*	*29*	*0*	
P	Florida	2	33	41	8	49
Q	Texas	2	38	42	4	46

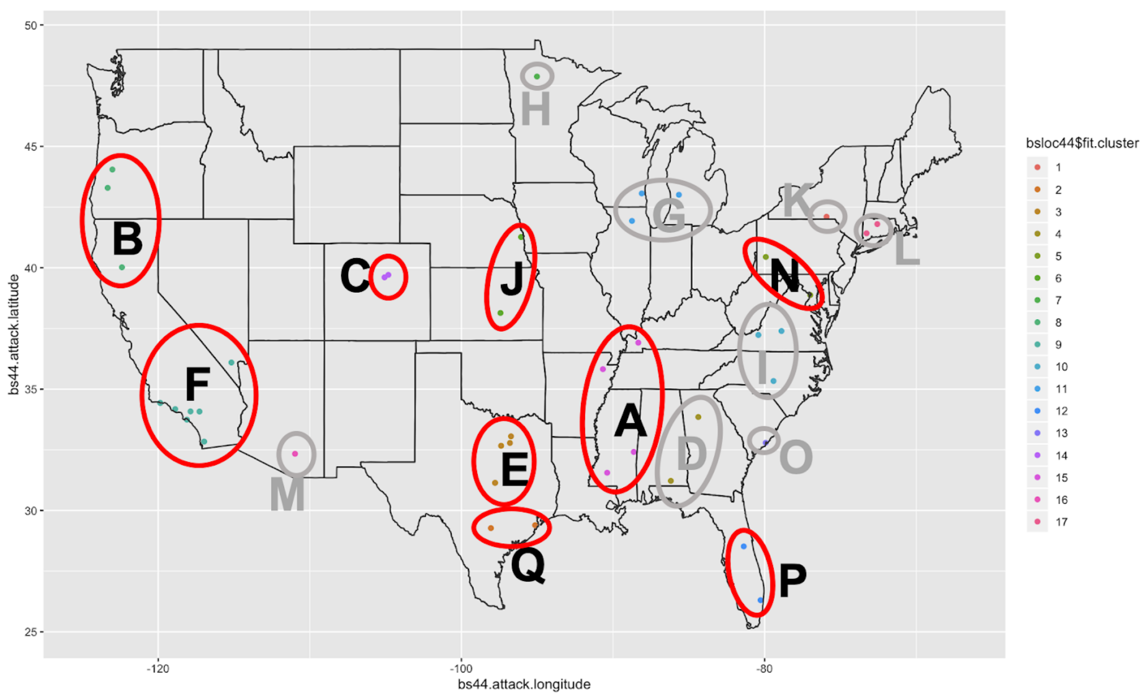


Figure 16 Black Swan Shooting highest risk hotspot map

Anecdotally, a quick check was conducted for new Black Swan Shootings outside the temporal parameters of the current study. As of March 2020, there had been five new events, occurring in Virginia, California, Texas twice, and Ohio. Each of these attacks falls near/within the 17 hotspots:

- Virginia Beach, VA: just south of high-risk hotspot N, east of hotspot I
- Gilroy, CA: in-between high-risk hotspots B and F
- El Paso, TX: east of hotspot M (a prior single-event hotspot)
- Dayton, OH: southeast of hotspot G, west of high-risk hotspot N
- Odessa, TX: west of high-risk hotspots E and Q, near the El Paso attack

Thus, these clustering patterns are remaining stable over time, indicative of spatial contagion effects.

Time

Temporally, Black Swan Shootings can initially be organized by year. On average, there are two attacks annually. However, as Figure 17 indicates, there is an increasing linear trend in attacks per year. After a relative lull during 2000-2004 of two total attacks, since 2005 twice has there been one attack in a year (2006 and 2013). Since 2005, the annual average has increased to 2.6 attacks; and since 2015, the annual average has jumped to 4.

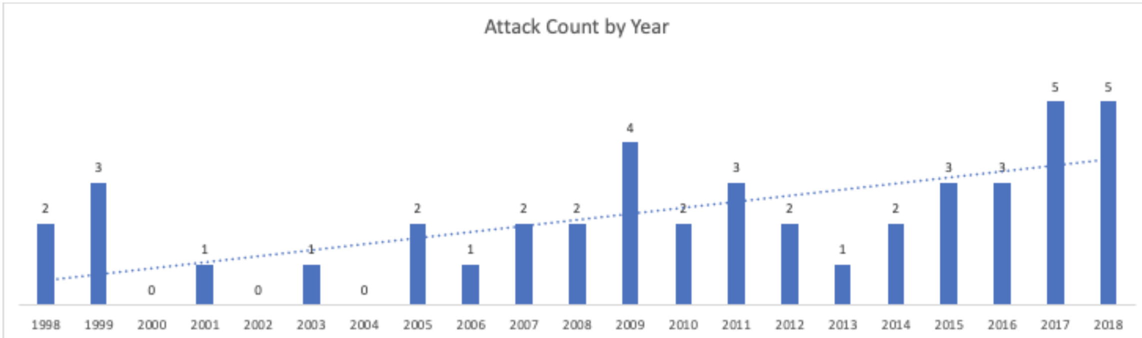


Figure 17 Black Swan Shootings per year

However, similar to counting by state, tabulating attacks by year is not ideal to identify true breaks and patterns. Thus, the 44 Black Swan attacks are analyzed for interval patterns. An interval is the number of days between attacks. This metric is calculated for the mean, median, standard deviation, and skew. The average interval is 175 days between attacks, with a standard deviation of 182 days, and a skewness of 1.84. Thus, these initial statistics give indications of spree activity, not significantly skewed intervals, and that 95% of the 44 events (approximately 42 events) have an interval of one year or less. This means that most of the time the next attack occurs within one year of the prior attack. Figure 18 shows a graph of attack intervals, and the significantly decreasing intervals over time. The last nine attacks have occurred on intervals below the mean. This is consistent with the Follman et al. (2018) and Krouse and Richardson (2015) findings indicating steadily increasing attack tempos since the mid-2000s.

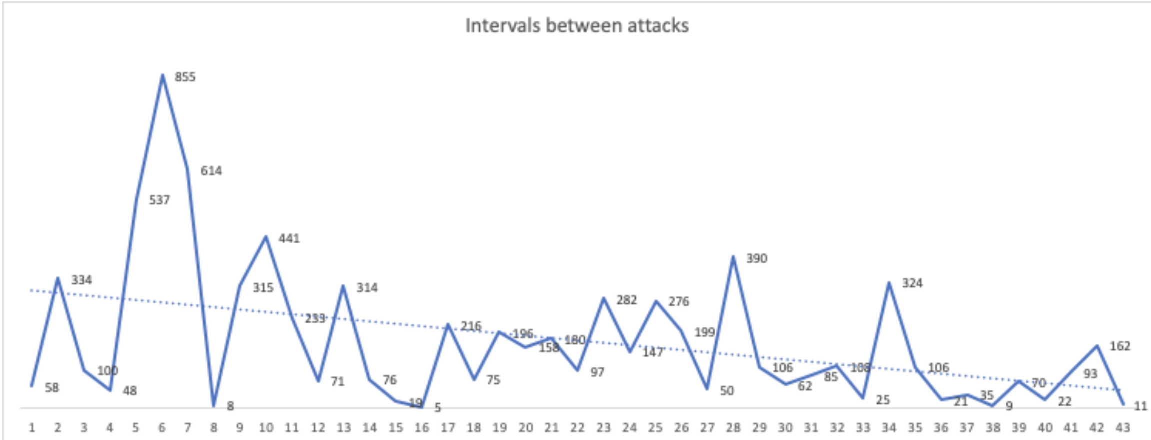


Figure 18 Black Swan Shootings intervals

Intervals can be used to calculate sprees and cycles of activity. Sprees and cycles are determined by analyzing distribution patterns of the interval calculations and identifying natural breaks. Sprees and cycles are the temporal equivalent of spatial clusters and hotspots. Figure 19 shows the interval values multiple ways. First, the top graph shows them chronologically. This demonstrates the overall trend of shrinking intervals over time. The second graph sorts the intervals in ascending order. This is done to identify clusters of similar values. The third graph is the calculated difference between the sorted intervals from the second graph. This graph makes it relatively easy to identify the biggest changes across intervals. Calculating the earliest natural breaks, using the second and third graphs, identifies the sprees and clusters. The shortest significant break occurs between interval values of 11 and 19. The next significant break occurs between intervals 25 and 35. The third significant cutoff happens at intervals 108 and 147 days. Thus, "primary" sprees are defined as occurring between 0-11 days after an initial attack; or the equivalent of 1-2 weeks. "Secondary" sprees occur between 0-35 days after an initial attack; or approximately one month. Cycles occur on 108-day rotations; which is roughly every 3.5 months. These temporal indicators can be re-coded to the event data and determine insights into the timing of events.

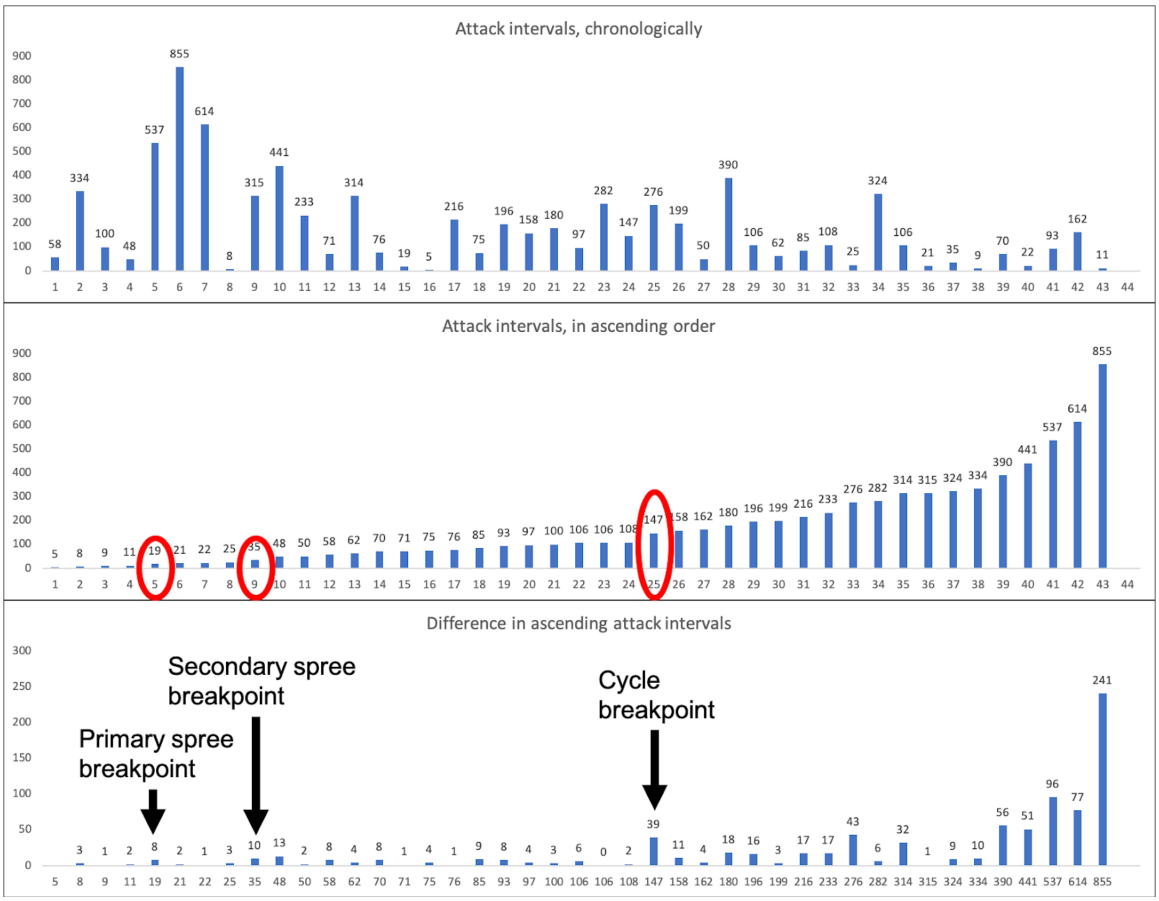


Figure 19 Black Swan Shootings intervals chronologically, in ascending order, and by calculated difference

Since cycles consist of events with intervals of 108 days or less between attacks, there are 20 cycles in this dataset. On average, these 20 cycles consist of 2.2 attacks over a 69-day duration, separated by 325 days of inactivity. Nine of the 20 cycles have one event, and four of 20 have three or more attacks. Table 13 provides details on each cycle. Similar to hotspots, this begs the question: can single events be a cycle? Yes, for the same reasons single events could be hotspots. Temporally, cycles are describing activity that follows a consistent, repeatable pattern. While the number of events and duration a cycle lasts will differ, there are regular breaks in between periods of activity, even for the single events. Events per cycle are increasing over time, with two of the three most recent cycles having the highest event counts. Since the cycles themselves are not increasing, but the number of events is, this is indicative of an increased tempo over time and potential contagion effects.

Table 13 Black Swan Shooting cycles

Cycle	Event count	Duration	Interval between cycles
1	2	59	n/a
2	3	149	334
3	1	1	537
4	1	1	855
5	2	9	614
6	1	1	315
7	1	1	441
8	2	72	233
9	4	101	314
10	2	76	216
11	1	1	196
12	1	1	158
13	2	98	180
14	1	1	282
15	1	1	147
16	1	1	276
17	2	51	199
18	6	387	390
19	8	357	324
20	2	12	162
Average	**2.2**	**69.0**	**324.9**
Average (non-Single Events)	**3.2**	**124.6**	**324.9**
Average (Cycles 6-20)	**2.3**	**77.4**	**255.5**

As Table 13 indicates, removing the nine single-attack cycles, the remaining 11 cycles have an average duration of 125 days, with an average of 3.18 attacks per cycle. Cycles can be so long because there is no natural break (defined here as an interval greater than 108 days) separating attacks. Over time, the interval between cycles is decreasing. Three of the four earliest cycles had the longest intervals. Examining the cycles over time, the first five cycles appear to have be unique intervals. Recalculating the average event count, duration, and interval for cycles 6-20 shows higher event counts, longer durations, and shorter cycle intervals.

Shifting to sprees, there are 40 primary sprees across the 44 attacks. The number of primary sprees is so high because they occur within 11 days or less from the prior attack, and this has happened four times: attacks 8 and 9 in March 2005; attacks 16 and 17 in March/April 2009; attacks 38 and 39 in November 2017; and attacks 43 and 44 in October/November 2018. Thus, 36 of the 40 primary sprees are single events. The most valuable finding derived from primary spree analysis is that two of them occurred across the seven most recent attacks. Aside from that, such a high number of single-event sprees offers little analytic utility for understanding copycat and contagion behavior; primary sprees will be ignored in favor of secondary sprees (now referred to as 'sprees'). Sprees are attack windows of 35 days or less. There are 35 sprees across the 44 attacks. This is still a high number, as 29 sprees are still single-event occurrences. Overall, sprees average 1.3 attacks, last for 5.4 days, and have an interval between attacks of 217 days. However, among the six multi-event sprees, the average number of attacks is 2.5 and the average duration is 27 days. Ten of the 12 most recent attacks occur in four different sprees, again suggesting that attacks have occurred closer together over time. Prior to this activity, there were two multi-event sprees. Table 14 details the cycle and spree statistics, and Figure 20 visualizes the cycles and sprees over time. Figure 20 demonstrates the temporal contagion effect; the longer the cycle, the more likely for sprees to occur within the cycle. Further, multi-event cycles are typically followed by multi-event cycles; all attacks since Attack 27 on 4/3/2014 have been part of a sequence of multiple events in a relatively short duration.

Table 14 Black Swan Shooting cycles and sprees

	Cycle				Spree		
#	Events	Duration	Interval	#	Events	Duration	Interval
1	2	59	n/a	1	1	1	n/a
				2	1	1	58
2	3	149	334	3	1	1	334
				4	1	1	100
				5	1	1	48
3	1	1	537	6	1	1	537
4	1	1	855	7	1	1	855
5	2	9	614	8	2	9	614
6	1	1	315	9	1	1	315
7	1	1	441	10	1	1	441
8	2	72	233	11	1	1	233
				12	1	1	71
9	4	101	314	13	1	1	314
				14	3	25	76
10	2	76	216	15	1	1	216
				16	1	1	75
11	1	1	196	17	1	1	196
12	1	1	158	18	1	1	158
13	2	98	180	19	1	1	180
				20	1	1	97
14	1	1	282	21	1	1	282
15	1	1	147	22	1	1	147
16	1	1	276	23	1	1	276
17	2	51	199	24	1	1	199
				25	1	1	50
18	6	387	390	26	1	1	390
				27	1	1	106
				28	1	1	62
				29	1	1	85
				30	2	26	108
19	8	357	324	31	1	1	324
				32	4	66	106
				33	2	23	70
				34	1	1	93
20	2	12	162	35	2	12	162
Average	2.2	69.0	324.9		1.3	5.4	217.0
Average (non-Singles)	3.3	136.2	261.3		2.5	26.8	189.3

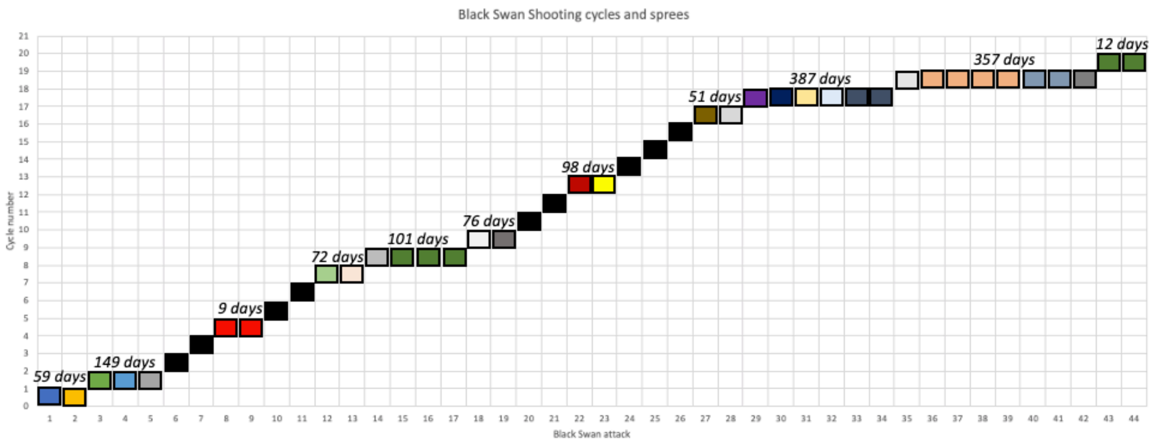


Figure 20 Black Swan Shootings cycles and sprees, with durations, over time

Alternatively, another way to examine attack timing is through probabilities. Table 15 highlights the likelihood of a new attack occurring across fixed calendar intervals. Of the 44 attacks, none have occurred on consecutive days, one has occurred within one week, eight have occurred within one month, 17 have occurred within three months, and 38 have occurred within one year of the prior attack. Thus, a majority of the time a new attack is expected within one year of the previous one, and often (but less than 50% of the time) a new attack occurs within three months of the last attack.

Table 15 Black Swan Shooting attack probabilities

Next attack in...	Interval Count	Probability
1 day	0	0%
1 week	1	2%
1 month	8	19%
1 quarter	17	40%
1 year	38	88%

Overall, the 44 Black Swan Shootings occur across 20 cycles and 35 sprees of activity. While multi-event sprees are not the norm, when they do happen (and they are happening more often), the next attack is expected to occur one week to one month later. If that next attack does not occur within a month, expect an approximately seven-month lull before any new activity. If the next attack does not occur by then, the cycle has ended, and a new attack is not expected until 11-12 months after the initial one. Similar to the hotspot analysis, these temporal findings can be anecdotally applied to recent mass shootings outside of the scope of this study. As of March 2020, the aforementioned five new events occurred on:

- 5/31/2019: Virginia Beach, VA
- 7/28/2019: Gilroy, CA
- 8/3/2019: El Paso, TX
- 8/4/2019: Dayton, OH
- 8/31/2019: Odessa, TX

The last attack in this study occurred on 11/7/2018. With an interval of 205 days until the Virginia Beach attack, cycle 20 has ended. Cycle 21 would comprise all five of these new attacks, with a duration of 93 days. Cycle 21 would consist of three sprees - two single-event sprees (Virginia Beach and Gilroy), followed by a 3-event, 29-day spree. As of March 2020, it has been more than 200 days since the Odessa attack, and cycle 21 has ended. Using the interval mean (237 days) and standard deviation (75 days) of cycles 8-

20 (which demonstrate a clearly similar trend of intervals), Cycle 22 would be expected to start by approximately February 2020 (so, now), and no later than July 2020.

Finally, combining the spatial and temporal findings, attacks in the same spree never occur in the same hotspot. In one cycle has attacks occurred in the same hotspot. During cycle 19 (eight attacks) there were two attacks in each hotspots A and Q; these are literally the first and last hotspots created. Further, there are eight total repeat counties described in the spatial analysis section - three for attacks, and three for residences. However, as Table 16 demonstrates, considering two of the attacks have the same residence and attack county, there are really four unique counties that experience repeat Black Swan Shooting activity: Arapahoe, Colorado; Bell, Texas; Santa Barbara, California; and Tarrant, Texas. The activity in Arapahoe is in a cluster by itself (Hotspot C), while the other three are part of larger multi-event clusters. These repeat locations include some of the most notorious attacks in the dataset, including Columbine High School (4/20/1999), the Dark Knight movie premier (7/20/2012), both Fort Hood attacks (11/5/2009, 4/3/2014), and the US Navy Yard in Washington, DC (9/15/1999). These attacks have above average victim counts, averaging nearly 30 total victims, with one attack having less than ten victims. On average, these repeat attacks occur 10 years apart, thus are never in the same spree or cycle.

Table 16 Black Swan Shooting repeat counties

Attack Date	Cycle	Spree	Hotspot	Total Victims	Years Between Attacks	Residence County	Residence State	Attack City	Attack County	Attack State
7/20/12	14	21	C	82	13.26	Arapahoe	Colorado	Aurora	Arapahoe	Colorado
4/20/99	2	3		37				Littleton		
4/3/14	17	24	E	15	4.41	Bell	Texas	Fort Hood	Bell	Texas
11/5/09	10	15		43						
1/30/06	6	9	F	8	8.32	Cibola	New Mexico	Goleta	Santa Barbara	California
5/23/14	17	25	B	19		Santa Barbara	California	Santa Barbara		
9/15/99	2	5	E	15	14.01	Tarrant	Texas	Fort Worth	Tarrant	Texas
9/16/13	16	23		20				Washington	Washington	District of Colombia
Average				*29.9*	*10.0*					

Overall, given this analysis, there is evidence supporting spatial and temporal contagion effects for Black Swan Shootings. As such, during the social factor variable selection, there will be a discussion on how these effects will be incorporated into the regression model.

Correlations matrix

Correlation matrices using the Pearson's R correlation score were generated for the 11 social factor variables for each of the count, rate, percentage change, and comparison to the nationwide mean metrics. For three of the metrics (count, rate, and percentage change), a correlation network was also generated. A correlation network weights the edges (the lines linking variables together) by their Pearson's R value, and clusters the nodes (the variables represented as points) to visualize the degree of connectivity across variables. The correlation network graph is a supplement visual aid for understanding the relationships between variables. These graphs help demonstrate variable clustering and distinctness. When examining a correlation matrix, it is difficult to manually interpret the relationships for more than one pair of variables. Thus, the network can highlight underlying relationships not readily visible. This will help later on for variable selection and modeling decisions.

First, Figure 21 is the correlation matrix for the count metric. The most notable finding is that nearly all of the correlations are positive. Additionally, the variables demonstrating the strongest correlations are largely socioeconomic and population-based counts. For example, population correlates strongly (0.9 and above) with race, sex, poverty, suicide, and drug and alcohol overdoses. Given the unnormalized nature of the count metric, this is somewhat expected.

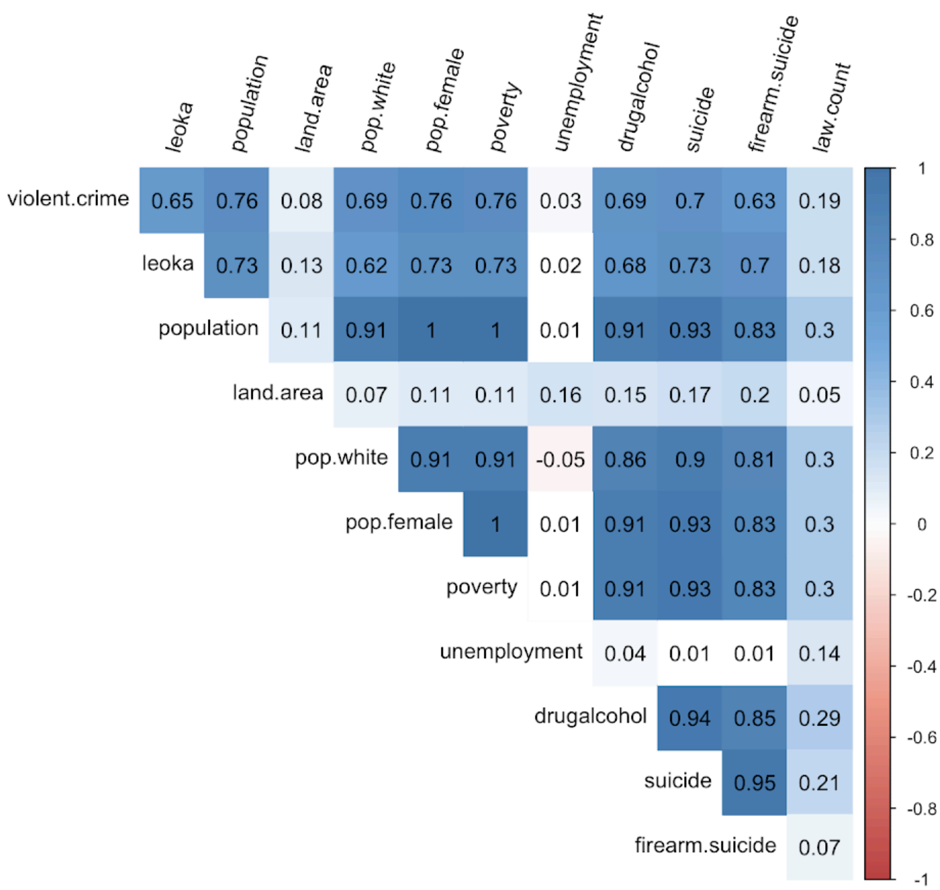


Figure 21 Correlation matrix for social factor counts

Further, Figure 22 demonstrates the largely anomalous relationships that county size ('land.area'), unemployment, and firearm laws have with the other eight variables.

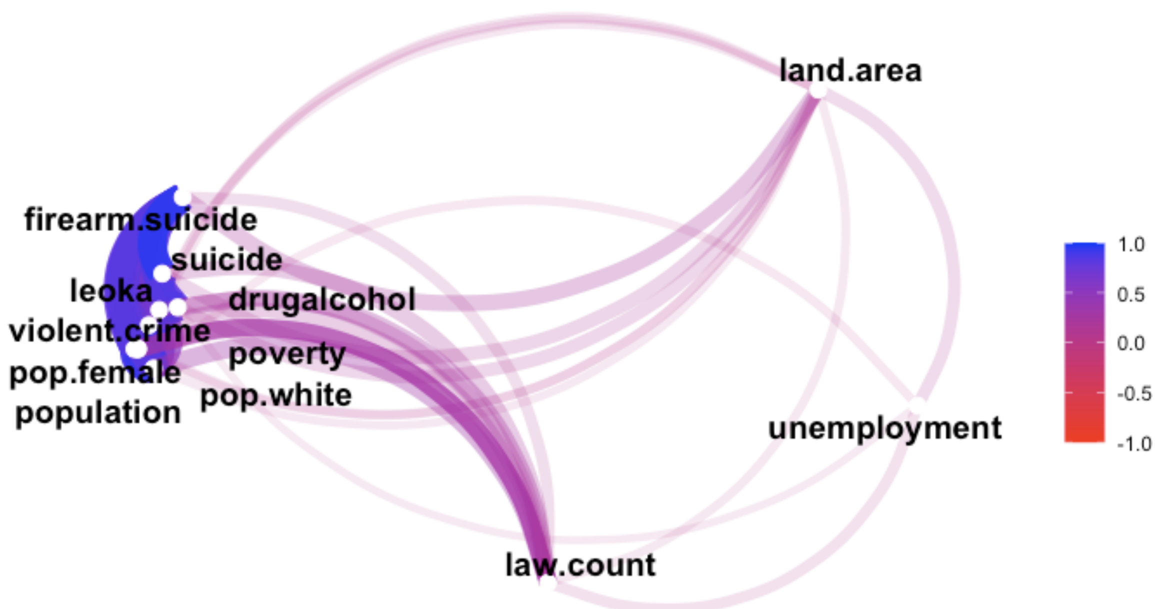


Figure 22 Correlation network for social factor counts

Second, Figure 23 displays the correlation matrix for the 11 social factors variables based on their rate calculations. This normalizes the count metric and offers a decidedly different view of these relationships. There is an overwhelmingly (and expected) strong positive relationship between suicide and percentage of suicide using a firearm (0.9). The next strongest positive correlations exist between drug and alcohol overdoses and suicide (0.54), violent crime and law enforcement officer killed or assaulted (LEOKA) (0.4), and drug and alcohol overdoses and percentage of suicide using a firearm (0.4). The strongest negative relationships exist between violent crime and race (-0.42), and sex and poverty (-0.31).

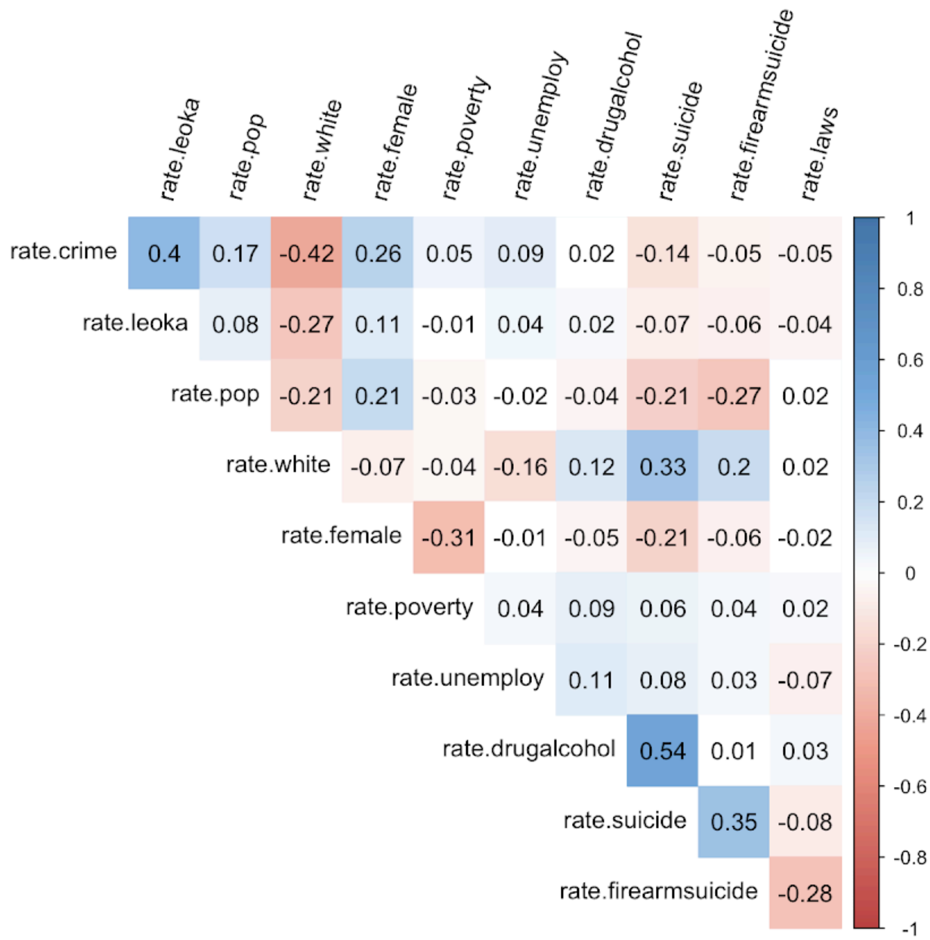


Figure 23 Correlation matrix for social factor rates

Next, Figure 24 displays the relationships of these rate calculations. Again, different than the count network, this chart clusters the 11 factors into roughly five groups: poverty and sex; violent crime and LEOKA; unemployment, race, and firearm laws; population; and suicide, drug and alcohol overdoses, and percent of suicides using a firearm. Once normalized, firearm laws and unemployment have stronger relationships, and population rate starts to look distinctly different than the other socioeconomic measures.

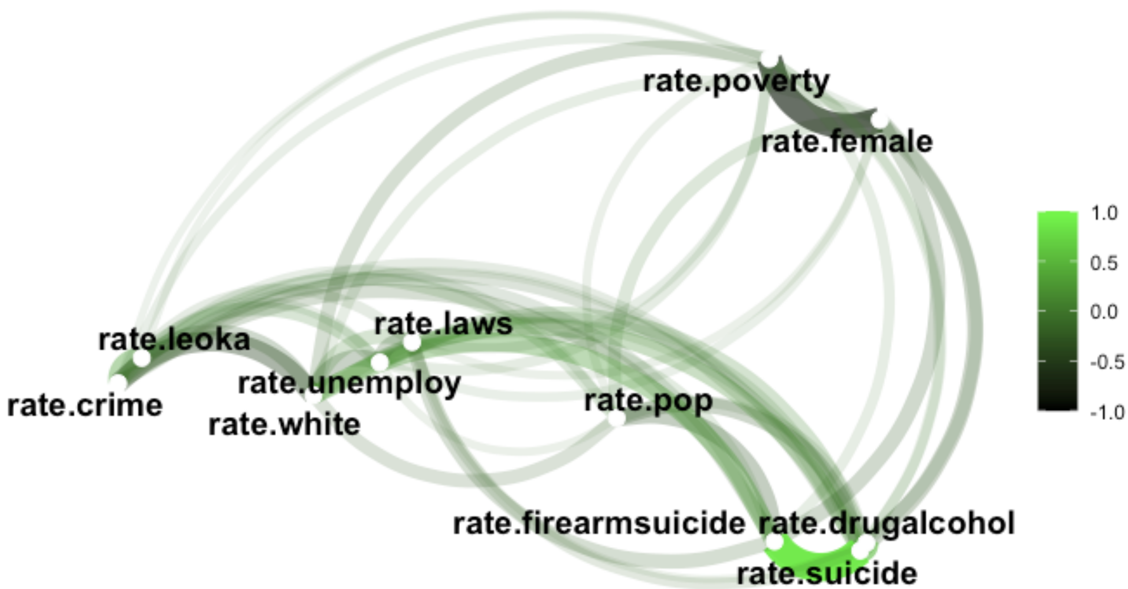


Figure 24 Correlation network for social factor rates

Third, Figure 25 is the correlation matrix for the annual percentage change. This matrix is significantly different from the matrices for count and rate. Nearly every pairwise comparison is close to zero, except for four: suicide and percentage of suicides with a firearm (0.66); race and sex (0.41), population density and race (-0.16), and suicide and drug and alcohol overdoses (0.14). Given that this matrix is for the annual change measurement, these relationships are expected. Interestingly, there are non-existent any relationships are for annual changes to crime and firearm laws. The LEOKA variable is not calculated for annual percentage change, because the annual counts and subsequent rates are so low.

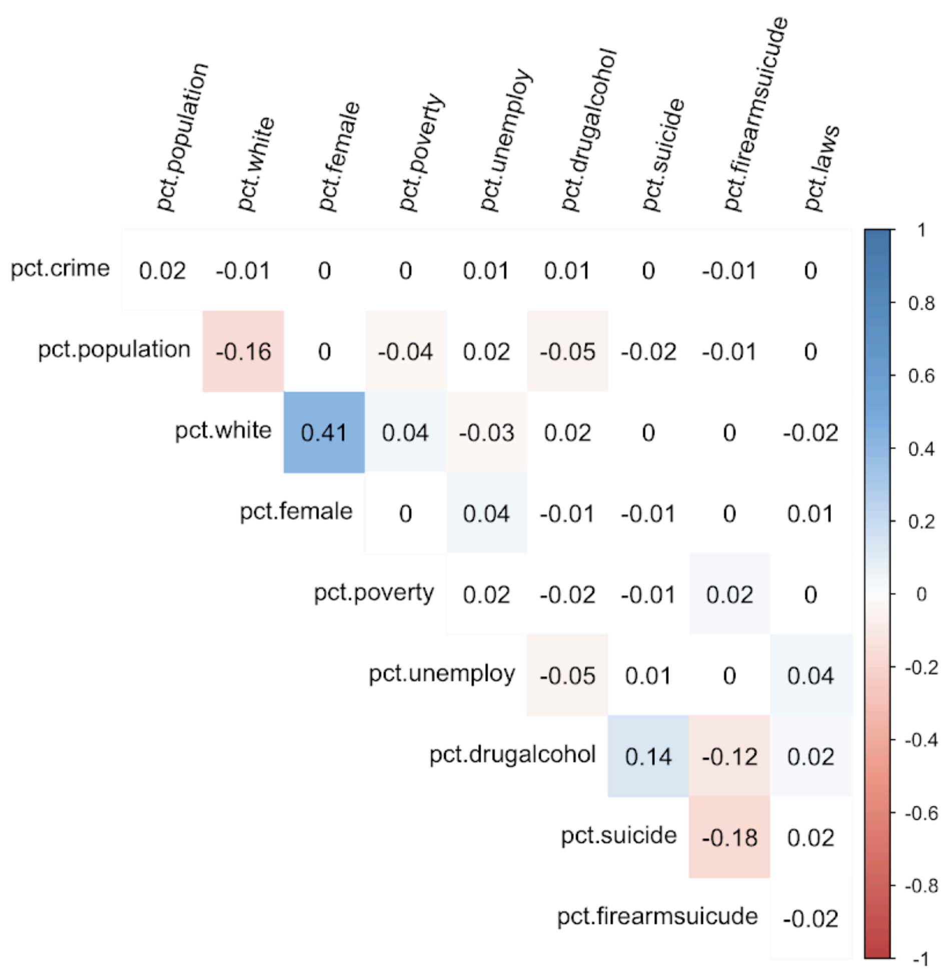


Figure 25 Correlation matrix for social factor annual percentage changes

Figure 26 renders the ten annual percentage change variables cleanly into three groups: race and sex; percent of suicide by firearm and suicide; and everything else. This is the clearest rendering across the different metrics of how these variables interact.

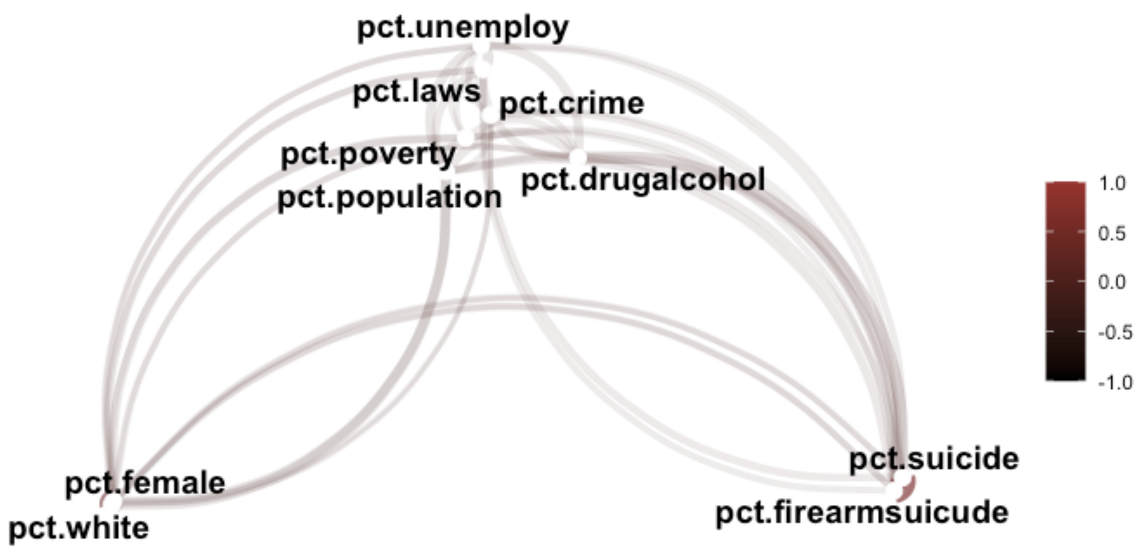


Figure 26 Correlation network for social factor annual percentage changes

Fourth, Figure 27 depicts the correlations for all 11 variables based on the comparison to the nationwide mean calculation. This closely resembles the rate matrix, which is expected. The comparison to the nationwide mean is more of an amplifying measure for rate, rather than its own independent calculation. This will be further addressed in the statistical tests section of this chapter. The strongest correlations here appear to be among socioeconomic variables with each other, and suicide, drug and alcohol overdoses, percent of firearm suicides. For firearm laws, these are the strongest correlations across any of the metrics; this includes the relationship between firearm laws and percent of firearm suicides (-0.37). This logically implies that more firearm laws correlate to less firearm suicides, and subsequently reduced firearm availability. This metric does not lend itself to a correlation network chart, so one was not created.

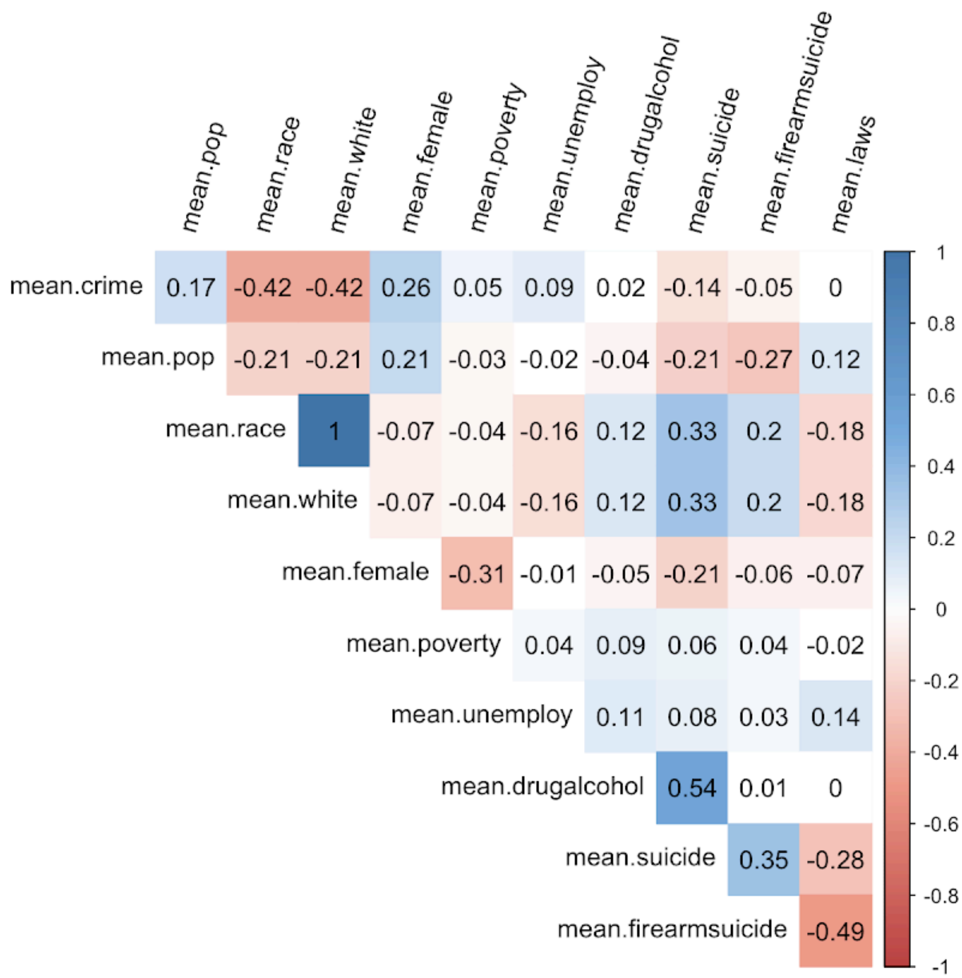


Figure 27 Correlation matrix for social factor comparison to the nationwide mean

Overall, these correlation matrices and networks highlight the degree to which the 11 variables are related. Among the four analytic groups these variables are operationalized from (violence, socioeconomics, mental health, and firearm preferences), there appear to be distinct, non-spurious relationships across the different metrics. For example, the annual percentage change for violent crime (from the violence layer) shows virtually no relationship with any other percentage change variable. For the socioeconomic layer, poverty rate and unemployment rate demonstrate no significant correlations to variables from the other three layers. For mental rate there is highly similar correlations within suicide and drug and alcohol overdoses, but weaker relationships with variables from the violence and socioeconomic layers. And for firearm preference calculations, there are distinct values for annual percentage change with the other three layers. Ultimately these measures will help improve variable selection for the regression model among potentially spurious relationships.

Statistical tests

For each social factor, a series of statistical tests are calculated. These tests include: a student's t-test for independent means; Cohen's d for effect size; and the Mann-Whitney U test (also known as the Wilcoxon rank-sum test) for non-parametric group comparison. Unless otherwise noted in the following sections, each test was calculated separately for attack and residence across the four metrics for each of the 11 variables, as described in Table 6. The scores for rate and comparison to the nationwide mean are merged as a single value, as the statistical tests yield the same results; thus, the rationale for comparison to the nationwide mean being an amplifying descriptor of rate. Four subsets were created from the overall panel time series database: Black Swan attack county-year observations (44 records); county-year observations without Black Swan attacks (66,085); Black Swan residence county-year observations (44 records); and county-year observations without Black Swan residences (66,085). As described in the event patterns section, the 44 records in each Black Swan dataset are a different compilation of observations, as the attack and residence counties are not the same. The following sections describe, by analytic layer and subsequent social factor variable, the results of each test for applicable metric for attack and residence measures.

Violence

The violence analytic layer is measured with violent crime and LEOKA. For violent crime, the results are mixed. For attack counties, the violent crime rate and annual percentage change are clearly different. As indicated in Table 17, counties experiencing Black Swan attacks have a more stable annual change (on average 72% change versus 85% for counties without Black Swan attacks) and higher violent crime rate (82 violent crimes per 100,000 people higher). The t-tests were not statistically significant, and the effect sizes were small, at best; however, the Mann-Whitney U was statistically significant (p-value = 0.0047). There were similar results for Black Swan residence counties. The violent crime rate is higher in Black Swan residence counties (309 per 100,000 versus 285 per 100,000), and the annual percentage change was smaller (63% versus 85%). The t-tests were not significant, the effect sizes were exceptionally small, and the Mann-Whitney U was significant for violent crime rate (p-value = 0.0185). These results suggest that violent crime rate is significantly higher for both Black Swan attack and residence counties when compared to counties without Black Swan activity.

Table 17 Statistical tests for violent crime

	Violent Crime	Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Rate**	367.44	284.90	0.0766	0.2691	0.0047
	Comparison to nationwide mean	82.49	-0.06			
	Annual percentage change	71.65	84.94	0.8047	0.0024	0.0930
Residence	Rate*	309.43	284.93	0.3449	0.0975	0.0185
	Comparison to nationwide mean	24.48	-0.02			
	Annual percentage change	62.90	84.94	0.6865	0.0041	0.4154

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

For LEOKA, the results are clearer and more consistent. For attack counties, both the count and rate are significantly higher. The count for LEOKA in Black Swan attack counties is more than ten times the count for LEOKA in counties without Black Swan attacks (177 versus 16, respectively). The t-test and Mann-Whitney U are statistically significant at the $p < 0.05$ and $p < 0.001$ levels, respectively. The Cohen's d metric shows moderately large effect sizes (0.49). LEOKA rate is only statistically significant for Mann-Whitney U (p-value = 0.0002). For residence counties, the results are similar. The count is ten times larger for LEOKA in Black Swan residence counties, with moderately large effect sizes (0.42), and Mann-Whitney U statistical significance at the $p < 0.01$ level. Table 18 displays the results.

Table 18 Statistical tests for LEOKA

LEOKA		Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Count***	176.64	15.62	0.0242	0.4887	0.0000
	Rate***	12.59	8.99	0.1319	0.1470	0.0002
Residence	Count***	152.05	15.63	0.0510	0.4196	0.0000
	Rate**	9.88	8.99	0.6316	0.0376	0.0088

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Socioeconomics

The second analytic layer is socioeconomics, which consists of five social factor variables: population, race, sex, poverty, and unemployment. Population is measured in two ways: the overall population (as a count), and the population density. These are treated as separate measures; population is a proxy for the size of the community, while population density is a normalized number for the congestion and potential criminal opportunities in an area. Population count demonstrates statistically significant differences for both Black Swan attack and residence counties across t-tests, Cohen's and Mann-Whitney U. In particular, the effect sizes for population are some of the highest observed for all social factors analyzed in this study. Counties experiencing Black Swan activity are nearly ten times more populated than counties without activity. Population density calculations indicate statistically significant relationships for both Black Swan attack and residence counties when compared to counties without such activity. As Table 19 indicates, the population density rate is four times higher in Black Swan attack counties (853 per 100,000 versus 209 per 100,000, respectively) and nearly three times higher for residence counties (621 per 100,000 versus 209 per 100,000, respectively). These results are statistically significant for both attack and residence with the t-test (at $p < 0.01$ for attack, $p < 0.001$ for residence) and Mann-Whitney U (at $p < 0.001$ for both attack and residence), with moderately large effect sizes (0.46 and 0.41 Cohen's d, respectively). For annual percentage change the results are similar: statistically significant differences for attack and residence counties, via t-tests (both at the $p < 0.001$ level) and Mann-Whitney U (both at the $p < 0.001$ level), with moderately large effect sizes (0.49 for attack counties, 0.55 for residence counties).

Table 19 Statistical tests for population

	Population Density	Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Count***	922496.90	95783.86	0.0012	0.7259	0.0000
	Rate***	852.89	209.14	0.0097	0.4623	0.0000
	Comparison to nationwide mean	643.33	-0.42			
	Annual percentage change***	-1.38	-0.38	0.0007	0.4944	0.0000
Residence	Count***	898778.00	95799.70	0.0018	0.6973	0.0000
	Rate***	620.54	209.29	0.0013	0.4051	0.0000
	Comparison to nationwide mean	410.97	-0.27			
	Annual percentage change***	-1.48	-0.38	0.0002	0.5520	0.0000

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Regarding race, calculations were run for 2000-2018, excluding the first two years of this study. US Census data prior to 2000 was found to be slightly inconsistent at the county level when compared to data from the 2000s. Race also exhibits strong relationships for both Black Swan attack and residence counties. As evidenced in Table 20, the Caucasian percentage of the population and the annual percentage change of the Caucasian percentage of the population are statistically significantly lower for Black Swan attack and residence counties. The results are nearly identical for attack and residence activity; statistical significance is found for percentage with t-tests (at the $p < 0.01$ level), and Mann-Whitney U ($p < 0.01$), and with exceptional high effect sizes

(Cohen's d values greater than 0.84). Annual percentage is statistically significant for attack and residence counties; however, the effect sizes are drastically lower.

Table 20 Statistical tests for race

Race		Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Percentage***	0.63	0.79	0.0000	0.8480	0.0000
	Comparison to nationwide mean	-0.16	0.00			
	Annual percentage change***	0.68	0.46	0.2148	0.1648	0.0097
Residence	Percentage***	0.62	0.79	0.0000	0.8834	0.0000
	Comparison to nationwide mean	-0.17	0.00			
	Annual percentage change***	0.77	0.46	0.0314	0.2460	0.0009

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Metrics for sex, measured as the percentage of the female population, were similarly compiled for 2000-2018. The percentage metric was found to be statistically significant for both attack and residence counties, via t-test ($p < 0.01$) and Mann-Whitney U (for attack counties, $p < 0.05$). For both sets, the effect sizes were low-moderate (0.36 for attack counties, 0.30 for residence counties). Annual percentage change showed no significant results across any test (see Table 21). Thus, the percentage of the female population is higher in counties with Black Swan activity.

Table 21 Statistical tests for sex

Sex		Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Percentage***	0.51	0.50	0.0011	0.3580	0.0324
	Comparison to nationwide mean	0.01	0.00			
	Annual percentage change	-0.23	0.07	0.1039	0.2699	0.1051
Residence	Percentage**	0.50	0.50	0.0027	0.3044	0.0909
	Comparison to nationwide mean	0.01	0.00			
	Annual percentage change	-0.16	0.07	0.1755	0.2149	0.1408

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Poverty is measured as a percentage of the county population. Interestingly, the poverty variables did not have any significant differences for attack or residence counties when compared to the rest of the country. As Table 22 displays, the percentage and annual percentage change for Black Swan counties were similar to non-Black Swan counties. Across the t-tests, Cohen's d, and Mann-Whitney tests, there were no significant differences or exceptional effect sizes. Poverty is not expected to be incorporated into any further calculations for Black Swan activity.

Table 22 Statistical tests for poverty

	Poverty	Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Percentage	2.42	2.99	0.2407	0.1386	0.9219
	Comparison to nationwide mean	-0.56	0.00			
	Annual percentage change	0.64	-8.66	0.6645	0.0045	0.3878
Residence	Percentage	2.55	2.99	0.4177	0.1028	0.6124
	Comparison to nationwide mean	-0.43	0.00			
	Annual percentage change	7.51	-8.66	0.3411	0.0078	0.6534

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

The final socioeconomic variable is unemployment. Unemployment is also measured as a percentage of the total county population. Similar to poverty, unemployment in Black Swan attack and residence counties shows no significant differences to counties without Black Swan attacks and residences. This is true for rate and annual percentage change, with no statistically significant differences in t-tests or Mann-Whitney, and exceptionally small effect sizes in Cohen's d. Table 23 shows the results. Unemployment is also not expected to be involved in any modeling for this study.

Table 23 Statistical tests for unemployment

	Unemployment	Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Percentage	5.79	6.02	0.4897	0.0924	0.8485
	Comparison to nationwide mean	-0.23	0.00			
	Annual percentage change	3.29	2.47	0.7263	0.0504	0.5637
Residence	Percentage	5.73	6.02	0.3643	0.1194	0.8602
	Comparison to nationwide mean	-0.29	0.00			
	Annual percentage change	2.85	2.47	0.8741	0.0229	0.7082

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Mental health

The third analytic layer is mental health. Designed to measure quality of life variables, mental health includes drug and alcohol overdoses, and suicides. For drug and alcohol overdoses, the rate for both attack and residence Black Swan counties was found to be significantly lower. Both t-tests were statistically significant at the $p < 0.05$ level, and the Mann-Whitney U was statistically significant for residence counties at the $p < 0.05$ level. In both cases the effect sizes were not low, but not quite moderate (0.37 for attack counties, 0.38 for residence counties). As Table 24 demonstrates, the annual percentage change statistics were not found to be significantly different from the rest of the country.

Table 24 Statistical tests for drug and alcohol overdoses

	Drug and Alcohol	Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Rate*	23.06	28.17	0.0101	0.3693	0.0555
	Comparison to nationwide mean	-5.09	0.01			
	Annual percentage change	-5.39	-0.68	0.1297	0.2058	0.6528
Residence	Rate*	22.74	28.17	0.0144	0.3776	0.0206
	Comparison to nationwide mean	-5.42	0.01			
	Annual percentage change	-5.37	-0.68	0.1897	0.1969	0.5259

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Suicide metrics were significantly different for one measure: the suicide rate for Black Swan attack counties, which was significantly lower than the rest of the country. This finding was statistically significant for the t-test ($p < 0.01$) and the Mann-Whitney U ($p < 0.05$), and the effect size was moderate (Cohen's d of 0.44). The other measures for suicide, including attack county annual percentage change, residence county rate, and residence county annual percentage change, had no significant differences to counties without Black Swan activity. See Table 25 for further details.

Table 25 Statistical tests for suicides

Suicide		Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Rate**	12.73	15.33	0.0049	0.4408	0.0114
	Comparison to nationwide mean	-2.59	0.01			
	Annual percentage change	-2.13	1.67	0.3268	0.1572	0.5899
Residence	Rate	13.51	15.33	0.0531	0.3025	0.0995
	Comparison to nationwide mean	-1.81	0.00			
	Annual percentage change	2.64	1.66	0.9156	0.0255	0.4517

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Firearm preferences

The fourth analytic layer is firearm preferences, which is measured as firearm availability and firearm political sentiment. Firearm availability is examined with a proxy, using the percentage of suicides committed with a firearm. The significant difference for this variable was Black Swan attack counties have a significantly lower percentage of suicides involving a firearm than counties without Black Swan attacks. This is statistically significant using t-tests and Mann-Whitney U (both at the $p < 0.05$ level), with a moderate effect size (Cohen's d of 0.36). Table 26 shows the remaining calculations have no significant difference from the rest of the country. Due to relatively poor data quality from CDC, annual percentage change was not calculated; too often there were not consecutive years of reported data for a given county for this variable.

Table 26 Statistical tests for percentage of suicides involving a firearm

	Firearm Suicide	Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Percentage*	0.51	0.56	0.0448	0.3610	0.0504
	Comparison to nationwide mean	-25.61	-25.56			
	Annual percentage change	n/a				
Residence	Percentage	0.54	0.56	0.3961	0.1468	0.4472
	Comparison to nationwide mean	-25.58	-25.56			
	Annual percentage change	n/a				

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Firearm political sentiment is expressed as the number of firearm laws in a state each year. When analyzing the state laws, 43 of the 44 attack site locations were able to be used. One of the attacks, 4/3/2014 at the U.S. Navy Yard, occurred in the District of Columbia; the State Firearm Law Database does not track D.C. firearm legislation counts. However, the offender residence for that attack is measured, so the full 44 residence counties are compared to state firearm legislative strength. Of the three metrics calculated for firearm laws - count, rate, and annual percentage change of count – the count metric showed significant differences. As Table 27 highlights, count was significantly higher for attack and residence counties. For attack, Black Swan counties average 37 firearm laws, and non-Black Swan attack counties averaged 21. This was statistically significant with a t-test and Mann-Whitney U, both at the $p < 0.01$ level, with a moderately high effect size (Cohen's d = 0.60). This is similar for residence: Black Swan counties average 34 firearm laws, counties without Black Swan residences average 21. This was statistically significant for t-tests and Mann-Whitney U ($p < 0.01$), with moderately high effect size (Cohen's d = 0.53). Rate and annual percentage change exhibit no significant differences for Black Swan activity.

Table 27 Statistical tests for state firearm laws

	Firearm Legislation	Black Swan Mean	Non-Black Swan Mean	t-test p-value	Cohen's d	Mann-Whitney U p-value
Attack	Count***	37.05	20.76	0.0028	0.5999	0.0010
	Comparison to nationwide mean	16.27	-0.01			
	Rate	0.36	0.42	0.3901	0.1284	0.0947
	Annual percentage change	1.23	1.77	0.8018	0.0293	0.1290
Residence	Count**	34.77	20.77	0.0065	0.5308	0.0032
	Comparison to nationwide mean	14.00	-0.01			
	Rate	0.36	0.42	0.3755	0.1305	0.0842
	Annual percentage change	1.20	1.77	0.7870	0.0310	0.1334

** denotes statistical significance at the p < 0.05 level; ** statistically significant at the p < 0.01 level; *** statistically significant at the p < 0.001 level*

Social factor variable selection

Variable selection is a three-step process. First, using the statistical calculations, any viable metrics for variables that demonstrate a significant difference for Black Swan activity are identified. A viable metric includes count, percentage, rate, and annual percentage change. As previously described, the comparison to the nationwide mean metrics are a descriptor and will help with understanding the direction and magnitude of the rate calculations. Thus, they are excluded from these lists and will be used to help in the next step. A significant difference is described as t-test or Mann-Whitney U test statistically significant at the $p < 0.05$ or lower, or a Cohen's d effect size at 0.30 or greater. Table 28 below highlights all of the selected variables, based on these filters, for attack and residence counties.

Table 28 Initial social factor variable selections

Type	Layer	Social factor variable	Metric	Significance		
				T Test	Cohen D	Mann-Whitney U
Attack	Violence	Violent Crime	Rate	x	x	0.005
		LEOKA	Count	0.024	0.489	0.000
			Rate	x	x	0.000
	Socioeconomics	Population	Count	0.001	0.726	0.000
		Population Density	Rate	0.009	0.462	0.000
			Annual percentage change	0.001	0.494	0.000
		Race	Percentage	0.000	0.848	0.000
			Annual percentage change change	x	x	0.010
		Sex	Percentage	0.001	0.358	0.032
	Mental health	Drug and Alcohol Overdoses	Rate	0.010	0.370	x
		Suicide	Rate	0.005	0.441	0.011
	Firearm Preferences	Firearm suicides	Percentage	0.045	0.361	0.050
		Firearm laws	Count	0.003	0.600	0.001
Residence	Violence	Violent Crime	Rate	x	x	0.019
		LEOKA	Count	0.051	0.420	0.000
			Rate	x	x	0.000
	Socioeconomics	Population	Count	0.002	0.697	0.000
		Population Density	Rate	0.001	0.405	0.000
			Annual percentage change	0.000	0.552	0.000
		Race	Percentage	0.000	0.883	0.000
			Annual percentage change change	0.031	x	0.001
		Sex	Percentage	0.003	0.304	x
	Mental health	Drug and Alcohol Overdoses	Rate	0.014	0.378	0.021
		Suicide	Rate	x	0.302	x
	Firearm Preferences	Firearm laws	Count	0.007	0.531	0.003

This process initially identifies 13 metrics across ten variables for Black Swan attack counties, and 12 metrics across nine variables for Black Swan residence counties.

Next, these remaining variables are compared to each other to identify the strongest, most salient relationships. When two similar variables are examined, the number of significant results, the magnitude of significance, the uniqueness from the correlation tests, and the interpretability of modeling are used to determine the most suitable variables for modeling. This process reduces the variable selection to one metric per social factor variable and removes any noise and potentially spurious correlations from the model.

There are multiple variables that are eliminated using this process. Violent crime rate and LEOKA rate have a 0.4 correlation (Figure 23). Violent crime rate and LEOKA rate are also clustered tightly on the correlation network graph for rates (Figure 24). Further, LEOKA count has no significant correlations with other variables, and the count metric for LEOKA is statistically significant based on t-tests and Mann-Whitney U at the $p < 0.05$ level. The count metric for LEOKA also has a moderately high effect size, with a Cohen's d of 0.49. However, despite the LEOKA count metric testing consistently better than the LEOKA rate, LEOKA is an artificially inflated number driven by the size of a community and the size of the police force. LEOKA rate makes more sense from a model interpretation standpoint and given its better performance and relationship to violence crime rates, LEOKA rate becomes part of the model, and LEOKA count is removed, for both attack and residence Black Swan activity. Violent crime rate and race percentage have a -0.42 correlation (Figure 23); however, these variables do not cluster in the network graph from Figure 24. For race, the percentage calculation performs better than the annual percentage change. For both attacks and residences, the race percentage is statistically significant from both tests and has exceptionally high effect sizes (above 0.84) - the highest in the entire dataset. The annual percentage change metric for race does not perform as well as the percentage; annual percentage change is removed from the variable selection. In the context of all the pairwise rate relationships, violent crime has relatively strong correlations with LEOKA, population density, race, sex, and suicide. However, violent crime rate measures statistically significant on Mann-Whitney and does not have a qualifying effect size. Given the lack of independence, and the stronger performances from LEOKA, population density, race, sex, and suicide, violent crime is removed from both attack and residence modeling.

A few other modeling decisions are determined from this process. Population density has significant results for rate and annual percentage change across attack and residence counties. However, the annual percentage change metric has consistently lower p-values and larger effect sizes, so rate is removed, and annual percentage change is kept for both the attack and residence models. Per Figure 23, drug and alcohol overdoses rates and suicide rates correlate relatively strongly (0.54), and closely align in the correlation network (Figure 24). For attack counties, suicide rate is statistically significant for both measures, and has a moderately high effect size (0.441). Drug and alcohol rate for attack counties is significant for a t-test, and at a higher p-value, with a smaller effect size (0.370). Thus, for the attack model, suicide rate is kept, drug and alcohol overdose rate are removed. For the residence model, drug and alcohol rate is statistically significant for both measures and has a slightly below moderate effect size (0.378). Suicide rate for residence counties is not statistically significant at all and has a lower effect size. Drug and alcohol overdose rate are kept for residence counties, and suicide rate is removed. In Figure 23, percentage of suicides committed with a firearm has a moderate correlation with suicide rate (0.35), which is the third strongest for all pairwise rate correlations. These two measures are also significantly clustered on the network correlation graph. For attack counties, both are statistically significant across both measures, and have qualifying effect sizes; however, suicide rate has lower p-values and a higher effect size. For the attack model, suicide rate is kept, and firearm suicides is removed. Finally, firearm law count, for both the attack and residence model, has no significant potentially spurious relationships with any other remaining variables. It is statistically significant and has moderately high Cohen's d scores, so firearm law count remains for both the attack and residence models.

The filtered list of remaining social factor variables is described in Table 29. For both Black Swan attack and residence counties, all four analytic layers are represented across seven variables. The difference is the mental health variable. For attack counties, it is suicide rate. For residence counties, it is drug and alcohol overdoses.

Table 29 Analytically derived social factor variable selections

Type	Layer	Social factor variable	Metric	Significance		
				T Test	Cohen D	Mann-Whitney U
Attack	Violence	LEOKA	Rate	x	x	0.000
	Socioeconomics	Population	Count	0.001	0.726	0.000
		Population Density	Annual percentage change	0.001	0.494	0.000
		Race	Percentage	0.000	0.848	0.000
		Sex	Percentage	0.001	0.358	0.032
	Mental health	Suicide	Rate	0.005	0.441	0.011
	Firearm preferences	Firearm laws	Count	0.003	0.600	0.001
Residence	Violence	LEOKA	Rate	x	x	0.000
	Socioeconomics	Population	Count	0.002	0.697	0.000
		Population Density	Annual percentage change	0.000	0.552	0.000
		Race	Percentage	0.000	0.883	0.000
		Sex	Percentage	0.003	0.304	x
	Mental health	Drug and Alcohol Overdoses	Rate	0.014	0.378	0.021
	Firearm preferences	Firearm laws	Count	0.007	0.531	0.003

Third, there are spatial and temporal contagion effects to capture in this model. Despite the variables in Table 29 being location-based, analysis earlier in this chapter identified unique clustering among the 44 Black Swan Shootings - so unique that it is not inherently captured in the aforementioned social factors. Thus, an additional indicator is added to the panel time series dataset for Black Swan hotspots. Any county-year observation within one of the 17 hotspots, in the year of the first attack for that hotspot or later, is flagged with a 1. All other counties are marked with a 0. Based on the mean nearest neighbor for all 44 attacks (184 kilometers/114 miles), as well as an examination of the county borders surrounding attack locations, all adjacent counties to an attack county are considered to be 'within' a hotspot, the mean nearest neighbor is calculated across the 44 attacks and applied to find the nearby counties. This process flags a total of 226 unique counties, and 2,387 new county-year observations. This effect is applied to the attack county regression model, since the spatial contagion was studied for attack locations and not residences. Further, there is a temporal contagion effect to add to the dataset. The analysis earlier in this chapter also identified cyclical activity, indicating that 88% of the time a new Black Swan Shooting occurs within one year of the prior attack (Table 15). Thus, starting in 1999, the number of attacks in the prior year is transformed into a lagged count for each county. This measure enriches 53,414 observations in the panel time series dataset. Similar to the spatial contagion, this effect applies to the attack model, as it does not intuitively make sense to include for modeling residences.

Before proceeding to the logit model, it is worth describing why such a manual process has been employed. First, it accounts for the spatial and temporal pattern analysis conducted earlier in this chapter. Second, it is a blend of quantitative and qualitative measures. The statistical tests are important, but equally important is the classification and logic of the inferred findings. For example, using LEOKA rate versus count makes more intuitive and interpretable sense, despite a potentially stronger statistical relationship. A blended approach ensures the appropriate variables at the appropriate levels are captured, to maximize the utility of the model's results. Third, King and Zang (2001) describe the best methods for modeling low-count events. These include avoiding sampling and increasing the depth of variables. In the current study, no sampling is conducted - all attacks are measured against all US counties. Further, by employing three statistical measures against 3-4 metrics for each social factor variable, analytic depth and rigor is achieved. With this in mind, a logistic regression model on such a small set of events is possible.

Regression model

Using the variables identified for Black Swan attack and residence counties, two logistic regression models are created. First, the Black Swan attack county model involves nine variables - seven social factors, and the two contagion effects. The results are shown in Table 30.

Table 30 Logistic regression model for Black Swan Shooting attack counties

Variable	Estimate	Std. Error	z value	p-value
Intercept	0.300467488	9.000047343	0.033385101	0.973367491
LEOKA rate	0.010045588	0.009253725	1.085572451	0.277668194
Population	0.000000158	0.000000119	1.327091389	0.184478453
Population density annual percent change	-0.175144482	0.186256648	-0.940339493	0.347043448
Race (Percentage of White)	-1.504891168	1.353190842	-1.11210564	0.266092732
Sex (Percentage of female)	-16.13132811	17.1446344	-0.940896594	0.346757853
Suicide rate	-0.003250453	0.047147958	-0.068941551	0.945036144
Firearm law count	0.00277107	0.008232096	0.336617812	0.736405027
Temporal contagion	-0.042241419	0.261730814	-0.161392608	0.871784188
Spatial contagion	4.455622302	0.779018734	5.719531643	0.000000011

From here, Equation 2 describes the regression equation is applied to the time series panel observations.

Equation 2 Black Swan attack county regression

County score = 0.300467488 + 0.010045588(LEOKA rate) + 0.000000158(Population) - 0.175144482(Population density annual percentage change) - 1.504891168(Percentage White) - 16.13132811(Percentage female) - 0.003250453(Suicide rate) + 0.00277107(Firearm law count) - 0.042241419(Temporal contagion) + 4.455622302(Spatial contagion)

Then, the probability for a Black Swan attack occurring in a given county is as follows:

Probability = 1/(1+exp(-score)) * 100

Of the 66,131 observations, this process is able to make predictions on 11,212 county-year observations. The reason for such a great discrepancy is the amount of missing data across some of the modeled variables; in particular, suicide rate. Suicide rate data from the CDC is inconsistently captured and recorded over time.

The logistic regression model for Black Swan residence counties is described is Table 31. This model involves seven social factors variables.

Table 31 Logistic regression model for Black Swan Shooting residence counties

Variable	Estimate	Std. Error	z value	p-value
Intercept	0.25843041	5.666814933	0.045604173	0.963625743
LEOKA rate	-0.019871888	0.015404661	-1.289991903	0.197053469
Population	0.000000354	0.000000099	3.569850803	0.000357185
Population density annual percent change	-0.23067285	0.082817937	-2.785300617	0.005347812
Race (Percentage of White)	-2.416346009	1.053563705	-2.293497771	0.021819364
Sex (Percentage of female)	-9.77769564	10.94122298	-0.893656555	0.371505677
Drug and Alcohol overdose rate	-0.024354858	0.017726675	-1.373910107	0.169469591
Firearm law count	0.008395655	0.007291259	1.151468394	0.249539591

Similar to the Black Swan attack county modeling, the process here involves scoring and assessing probability for all potential observations. The scoring equation, in Equation 3, is as follows:

Equation 3 Black Swan residence county regression

County score = 0.25843041 - 0.019871888(LEOKA rate) + 0.000000354(Population) - 0.23067285(Population density annual percentage change) - 2.416346009(Percentage White) - 9.77769564(Percentage female) - 0.024354858(Drug and alcohol overdose rate) + 0.008395655(Firearm law count)

Next, the probability equation is the same as the one for Black Swan attack counties:

Probability = 1/(1+exp(-score)) * 100

And similarly, of the 66,131 observations, this analysis makes predictions on 10,061 records. This is for the same reason as previously addressed - missing data on from CDC, in this case drug and alcohol overdose rates.

Model performance

For the Black Swan attack county model, there are 11,212 available predictions. The mean prediction is a 0.196% probability of an attack occurring in a given county for a given year, with a standard deviation of 0.764, and a skewness of 8.49. To identify the counties at highest risk in a given year, the predicted scores were filtered based on the mean and standard deviation. Any probability scoring above the 95% threshold of all predictions (the mean plus two standard deviations) was deemed to be 'high risk.' Of the initial 11,212 predicted values, this left 375 predicted attack counties for all time. The average probability for a high-risk county among these 375 predictions is 3.37%. Figure 28 highlights the number of predicted counties annually. The number of predicted counties each year ranges from 0 to 42, with an average of 21 counties at highest risk for an attack each year.

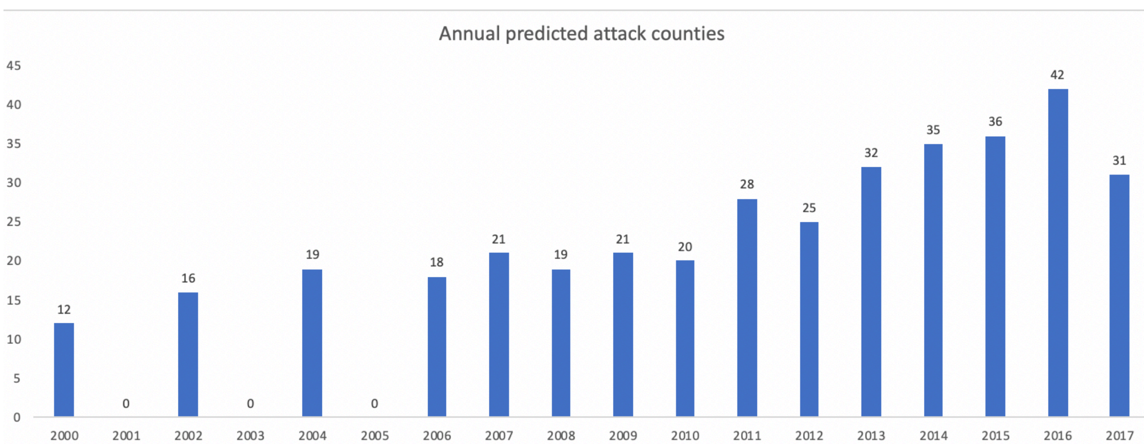


Figure 28 Black Swan annual highest risk attack counties

To assess model performance, spatial precision, spatial accuracy, and temporal accuracy are measured using the metric described in chapter three. As an example, Figure 29 demonstrates how the metric is calculated, using the 2006 high-risk attacks as an example. Each component of the metric is measured as one-third, with the overall score scaled 0-1. What is a “good” score? For the purposes of this study, any year with a metric score greater than 0.5 is considered analytically viable for two reasons. First, for a statistical model to be worthwhile and actionable, it should be better than guessing. Although somewhat arbitrary, the 0.5 threshold is deemed better than a coin flip. Second, based on the metric component weights, scoring exactly 50% across spatial precision, spatial accuracy, and temporal accuracy would yield a 0.5 metric score. Anything deemed analytically viable should perform than 50% overall across its components.

1. **For a single year** (2006, for example)…
    a. One attack county
    b. Eighteen high-risk counties
2. **Spatial precision**
    a. Percentage of country at high-risk
        a. 1 – (18 high-risk counties divided by 3,142 total U.S. counties) = **0.994**
3. **Spatial distance**
    a. Count, for all high-risk counties, of states separating prediction from actual attack
        a. 7x1 + 3x3 + 5x7 + 3x3 = 60
4. **Spatial accuracy**
    a. Proximity of high-risk counties are to the actual attack county
        a. 50 states divided spatial distance = **0.833**
5. **Temporal accuracy**
    a. How close the number of high-risk counties is to the number of attacks
        a. One attack divided by 18 high-risk counties = **0.056**
6. **Metric score**
    1. (1/3)(0.994) + (1/3)(0.833) + (1/3)(0.056) = **0.628**

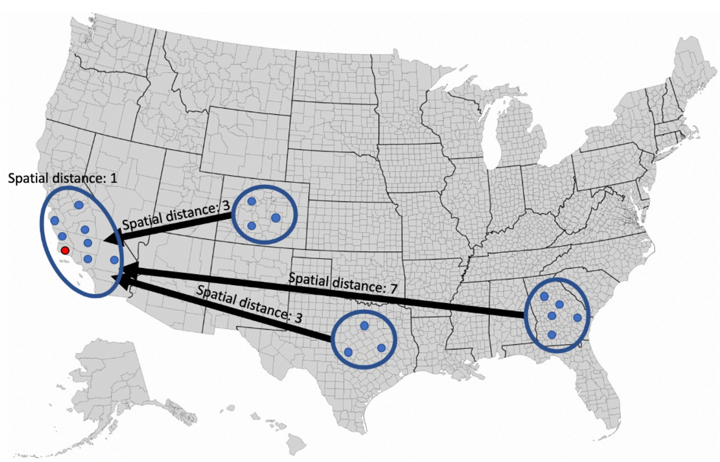


Figure 29 Black Swan metric calculation for high-risk attack counties in 2006

The results of these calculations are displayed in Table 32. When calculating spatial distance, the number of states separating the predicted county and attack county were counted. If the prediction was a bullseye (meaning directly predicting the county-year for an attack), the spatial distance is zero; if the state was correct but the attack was in a different county, the score was one.

Table 32 Black Swan attack county prediction performance

Year	Predicted counties	Attack Probability	Spatial precision	Spatial distance	Spatial accuracy	Temporal accuracy	Actual attack counties	Bullseye	State Hit	Metric score
2000	12	Average: 2.61%, Max: 3.58%, Min: 1.82%	0.996	0		0.000	0	0		0.332
2001	0				0.000	0.000	1	0		0.000
2002	16	Average: 2.86%, Max: 7.24%, Min: 1.93%	0.995	0		0.000	0	0		0.332
2003	0				0.000	0.000	1	0		0.000
2004	19	Average: 3.45%, Max: 9.12%, Min: 1.81%	0.994	0		0.000	0	0		0.331
2005	0				0.000	0.000	2	0		0.000
2006	18	Average: 3.51%, Max: 9.62%, Min: 1.76%	0.994	60	0.833	0.056	1	0	7	0.628
2007	21	Average: 3.35%, Max: 9.82%, Min: 1.74%	0.993	61	0.820	0.095	2	0	0	0.636
2008	19	Average: 3.84%, Max: 14.41%, Min: 1.75%	0.994	36	0.720	0.105	2	1	10	0.606
2009	21	Average: 4.25%, Max: 14.63%, Min: 1.74%	0.993	44	0.880	0.190	4	1	6	0.688
2010	20	Average: 3.38%, Max: 13.42%, Min: 1.87%	0.994	138	0.362	0.100	2	0	0	0.485
2011	28	Average: 3.50%, Max: 14.65%, Min: 1.75%	0.991	62	0.806	0.107	3	1	13	0.635
2012	25	Average: 3.43%, Max: 13.49%, Min: 1.74%	0.992	71	0.704	0.080	2	0	2	0.592
2013	32	Average: 3.51%, Max: 14.17%, Min: 1.73%	0.990	230	0.217	0.031	1	0	0	0.413
2014	35	Average: 3.50%, Max: 15.16%, Min: 1.74%	0.989	87	0.575	0.057	2	1	16	0.540
2015	36	Average: 3.50%, Max: 14.67%, Min: 1.77%	0.989	73	0.685	0.083	3	1	10	0.586
2016	42	Average: 3.21%, Max: 13.58%, Min: 1.73%	0.987	88	0.568	0.071	3	1	13	0.542
2017	31	Average: 2.62%, Max: 5.74%, Min: 1.73%	0.990	60	0.833	0.161	5	2	16	0.662
Sum	**375**	**n/a**	**n/a**	**1010**	**n/a**	**n/a**	**34**	**8**	**93**	**n/a**
Average	**20.83**	**Average: 3.37%, Max: 11.55%, Min: 1.77%**	**0.99**	**84.17**	**0.53**	**0.06**	**1.89**	**0.44**	**7.75**	**0.44**

Overall, the average metric score is below 50%, at 0.44. However, examining that overall number would be misleading. First, the first six years of the model perform poorly, followed by 12 years of better scores. This includes ten of the years since 2006 being analytically viable, exceeding the 0.5 threshold. Second, spatial precision has gotten marginally worse over time. However, any year is never lower than 98.7%, meaning that this model never deems more than 1.3% of the entire country to be at high risk for a Black Swan Shooting. That is an excellent barometer, given the rarity of events. Third, spatial distance, known as the number of states in between a high-risk county and an actual attack, varies year to year. Two of the years returning results, 2010 and 2013, were significantly poor performing. Five of the years (2006, 2007, 2009, 2011, and 2017) were significantly above average for spatial accuracy, at greater than 80%. Fourth, temporal accuracy hurts the overall performance. This is somewhat expected, given that the highest number of attacks in any year is five, and the lowest number of predicted counties is 12.

More noteworthy, this model successfully predicted the exact county and year for eight attacks. While 34 of 44 Black Swan Shootings occurred during the modeled time period (1998, 1999, and 2018 were not calculated), this is a significant finding. Further, all of these "bullseyes" occurred since 2008, including every year since 2014, when Black Swans have become more prevalent. Thus, this model has successfully retro-predicted 30% of the locations of Black Swan Shootings since 2008. Broader, this model successfully predicted the state of occurrence for 93 of the 375 high-risk counties. All of these have been since 2006, making the model successful 28% of the time. Any year with a single bullseye or ten or more state matches has an overall metric score of at least 0.54.

Given this relatively strong performance, it is worth assessing the diversity of annual county predictions. There are 68 unique counties labeled as high-risk 375 times. Table 33 shows the number of repeats. While 31% of unique counties were predicted for 9 or more years, 44% of the high-risk counties were predicted for two years or less. This demonstrates the model's ability to generate a relatively unique list of high-risk attack counties annually.

Table 33 Black Swan attack county prediction totals

Attack Counties

Years Predicted	County Count
15	4 (1 Bullseye)
14	2
13	3 (1 Bullseye)
11	4
10	2
9	6 (4 Bullseyes)
8	2
6	5
5	3
4	5 (1 Bullseye)
3	2
2	13 (1 Bullseye)
1	17
Total	***68***

This process is repeated for Black Swan residence counties. Of the 10,062 available predictions, the mean prediction was a 0.15% probability for a county in a given year, with a standard deviation of 0.463 and a skewness of 24.38. Filtering for the highest risk counties in this model uses the mean plus one standard deviation. Based on the 68% barometer, there are 232 viable predictions. Of these 232 predictions, the mean probability for a high-risk Black Swan residential county is 1.5%. By year, the high-risk county counts are displayed in Figure 30. The annual counts range from 4 to 26 predicted counties, with an average of 13 predicted per year.

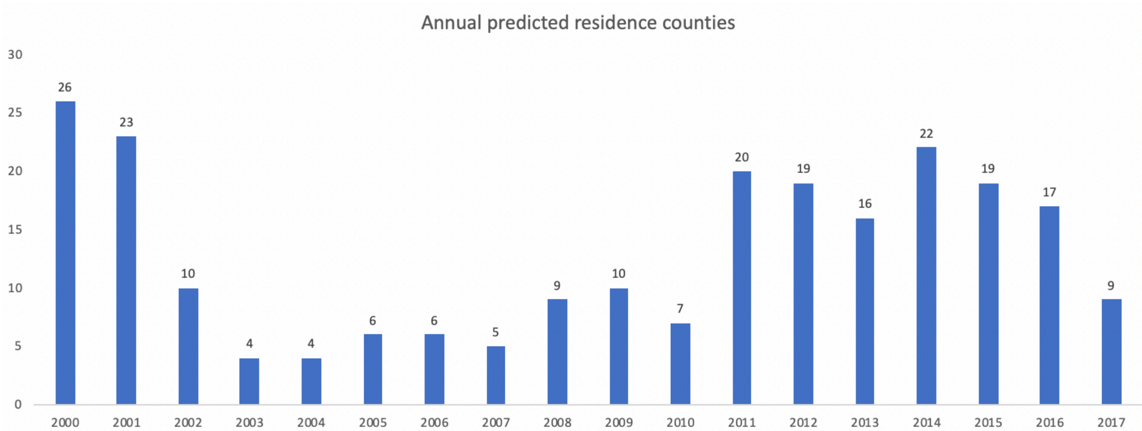


Figure 30 Black Swan annual highest risk residence counties

The metrics for annual residence prediction performance are displayed in Table 34. Overall, the metric score is higher for the residential model, at 56%. This model appears to be more consistent, but not as high performing as the attack model. There are 13 analytically viable years, starting in 2001 and including every year since 2009. Spatial precision for high-risk residential counties is never lower than 99.2% of the country. Spatial distance is nearly 50% less compared to the attack model. This indicates that the "near misses" are much closer to the actual residence. Spatial accuracy is actually lower in the residential model, despite the distance being so much closer. The cause for the difference may be that the residential model consistently misses the exact location (less bullseyes) but is in the right vicinity and much closer. Temporal accuracy performs better in the residential model, partially because of the lack of zero-prediction years. This is driven by not including the temporal contagion effects in this model. This model had six bullseyes, where the county and year of residence for a Black Swan Shooter was an exact match. Five of the six bullseyes have occurred since 2008. Also, there have been 69 of the 232 high-risk counties hitting in the actual state of residence. Since 2008, 40% of the high-risk counties have been in the state of the Black Swan residence. Given the high precision (never more than 22 counties in a given year since 2008 labeled as high-risk), this is a significant finding.

Table 34 Black Swan residence county prediction performance

Year	Predicted counties	Attack Probability	Spatial precision	Spatial distance	Spatial accuracy	Temporal accuracy	Actual residence counties	Bullseye	State Hit	Metric score
2000	26	Average: 1.26%, Max: 10.43%, Min: 0.62%	0.992		0.000	0.000	0	0	0	0.331
2001	23	Average: 1.40%, Max: 11.70%, Min: 0.61%	0.993	61	0.820	0.043	1	1	11	0.619
2002	10	Average: 2.19%, Max: 12.05%, Min: 0.65%	0.997		0.000	0.000	0	0	0	0.332
2003	4	Average: 0.99%, Max: 1.39%, Min: 0.68%	0.999	20	0.400	0.250	1	0	0	0.550
2004	4	Average: 1.06%, Max: 1.58%, Min: 0.65%	0.999		0.000	0.000	0	0	0	0.333
2005	6	Average: 0.89%, Max: 1.29%, Min: 0.64%	0.998	30	0.600	0.333	2	0	0	0.644
2006	6	Average: 0.94%, Max: 1.46%, Min: 0.62%	0.998	11	0.220	0.167	1	0	0	0.462
2007	5	Average: 0.90%, Max: 1.06%, Min: 0.67%	0.998	20	0.400	0.400	2	0	0	0.599
2008	9	Average: 2.08%, Max: 11.25%, Min: 0.65%	0.997	10	0.200	0.222	2	1	8	0.473
2009	10	Average: 1.88%, Max: 10.36%, Min: 0.63%	0.997	26	0.520	0.400	4	1	2	0.639
2010	7	Average: 2.52%, Max: 11.88%, Min: 0.82%	0.998	67	0.746	0.286	2	0	0	0.677
2011	20	Average: 1.55%, Max: 12.81%, Min: 0.62%	0.994	35	0.700	0.150	3	1	9	0.615
2012	19	Average: 1.70%, Max: 14.79%, Min: 0.63%	0.994	62	0.806	0.105	2	0	0	0.635
2013	16	Average: 1.93%, Max: 14.66%, Min: 0.64%	0.995	44	0.880	0.063	1	0	4	0.646
2014	22	Average: 1.60%, Max: 14.71%, Min: 0.63%	0.993	36	0.720	0.091	2	0	16	0.601
2015	19	Average: 1.58%, Max: 13.26%, Min: 0.62%	0.994	35	0.700	0.158	3	1	9	0.617
2016	17	Average: 1.62%, Max: 13.16%, Min: 0.63%	0.995	37	0.740	0.176	3	1	5	0.637
2017	9	Average: 0.85%, Max: 1.35%, Min: 0.61%	0.997	14	0.280	0.556	5	0	5	0.611
Sum	232	n/a	n/a	508.000	n/a	n/a	34	6	69	n/a
Average	12.89	Average: 1.50%, Max: 8.84%, Min: 0.65%	1.00	33.87	0.49	0.19	1.89	0.33	3.83	0.56

Similar to the attack county predictions, the residential model predicts a diverse set of counties as high-risk each year. There are 46 unique counties across the 232 high-risk predictions. As Table 35 demonstrates, while 28% of the counties are high-risk nine or more years (including five bullseyes), 57% of the counties are high-risk for two or less years (with one bullseye).

Table 35 Black Swan residence county prediction totals

Residence Counties

Years Predicted	County Count
17	4 (2 Bullseyes)
12	2 (2 Bullseyes)
10	1 (1 Bullseye)
9	6
7	3
5	3
4	1
2	10
1	16 (1 Bullseye)
Total	***46***

In conclusion, consider the nuance to each of these scores. Overall, this analysis demonstrates that location-based social factors appear to have at minimum a modest relationship with Black Swan Shootings, but more likely a strong relationship. In particular, for examining Black Swan attack counties, it appears that the relationship between these shootings and prior attacks via spatial and temporal contagion effects, LEOKA rate, county population, the annual percentage change of the population density, the percentage of the white population, the percentage of the female population, the suicide rate, and the number of firearm laws have the ability to identify a significant low number of high-risk counties in a given year. Similarly, when examining the residential county for a Black Swan Shooting, the LEOKA rate, county population, the annual percentage change of the population density, the percentage of white residents, the percentage of female residents, the drug and alcohol overdose rate, and the number of state firearm laws are able to identify a significantly low number of high-risk counties in a given year. The interpretations and implications for these findings will be discussed in the next chapter.

DISCUSSION AND CONCLUSIONS

Discussion

In this study, a Black Swan Shooting is defined as an event occurring in the United States since 1998 involving one or more coordinated perpetrators acting on behalf of themselves or a non-state-sponsored entity using one or more firearms to either kill eight or more people, injure 13 or people, or cause 15 or more total casualties. There are 44 such attacks. Prior research demonstrates mass violence is increasing in the U.S. at an alarming rate, and tends to occur in communities that are less social, less rural, have high rates of overcrowding, high incomes inequalities, greater firearm availability, and more permissive laws for firearm ownership, access, prevention, and safety (Markowiak et al., 2018; Webster et al., 2016; Reeping et al., 2019). Prior research on person-centric theories for mass violence causality remain largely untested and unproven for intervention, disruption, and deterrence of such attacks, given the limited current research. This study has taken a different approach, focusing on the location-based social factors identified in prior mass violence research to have a relationship to such attacks to develop a nationwide threat assessment for communities at highest risk for Black Swan Shootings to occur.

This study had three goals. First, this study sought to explain any attack contagion measures associated with Black Swan Shootings. The goal was to identify whether these attacks cluster spatially and occur on any distinct timelines, and if those patterns have predictive power. That answer, to a degree, is yes. Spatially, the 44 attacks occur in 17 distinct areas. Four of the 17 clusters have a single event, and each of those events are ten years or older. There are 26 states that have never had activity. Accounting for fallow clusters, there are nine currently active areas in the U.S. Temporally, Black Swan Shootings occur in repeatable cycles. When a new attack happens, there is a heightened window for another attack in the week to month that follows. After that, there is typically a seven-month lull. If there is still no new attack, the cycle has ended, and nothing occurs until nearly one year after the initial attack. A majority of attacks occur within one year of the previous one (88%), and in an existing hotspot. This degree of predictability over space and time is potentially stronger than previous studies have found. These contagion measures have the potential to serve as part of an alerting system for law enforcement, education, and health practitioner responses. In the immediate one-month aftermath of Black Swan Shooting, counties within nationwide hotspots could be put on high alert to focus attention on identifying person-centric early warning signs. Such a plan would leverage the findings of this study, combined with the prior research on behaviors and characteristics of mass violence perpetrators.

Second, this study has examined the relationship between Black Swan Shootings and location-based social factors. There are five analytic layers of social factors: the aforementioned attack contagion, violence, socioeconomics, mental health, and firearm preferences. These layers and their associated social factors were identified through known composite and community measures for mass violence. Each factor has already been identified as having a relationship to mass violence or has been derived through analysis of the common traits and behaviors of known mass violence perpetrators and extrapolated to the community level. Violence is measured with violent crime and law enforcement officers killed or assaulted in the line of duty. Socioeconomics includes measures for population, race, sex, poverty, and unemployment. Mental health measures prosocial activity, which in this case are substance abuse and quality of life, captured here as drug and alcohol overdoses and suicide. Firearm preferences are identified as firearm availability and firearm legislation. Firearm availability is measured using the percentage of suicides involving a firearm, and firearm legislation is the number of state firearm laws limiting permissive behaviors.

These 11 social factors were compared to counties of Black Swan attacks, and the counties of Black Swan residences. Here, the answer also appears to be yes: location-based social factors have unique relationships for communities experiencing Black Swan Shootings. Ultimately, Black Swan Shooting attack counties experience higher levels of LEOKA violence, have a dense population subject to influx, are more racially diverse, have more female residents, have lower suicide rates, and have nearly twice as many firearm laws, when compared to counties where Black Swan Shootings have not occurred. For Black Swan Shooting residence counties, LEOKA violence is ten times higher, populations are three times denser, more racially diverse, have more female residents, experience lower rates of drug and alcohol overdoses, and have more restrictive firearm laws compared to the rest of the country. For both Black Swan attack and residence counties, poverty and unemployment showed no significant relationships.

Similarly, violent crime rates and percentage of suicides committed by firearms had weak and potentially spurious relationships. Thus, the combination of these social factors paints a distinct portrait of a community. The seemingly archetype county at high-risk for Black Swan Shooting activity is large, dense, race and gender diverse, at odds with law enforcement, relatively safe and not prone to substance abuse, and strict on firearms. This is consistent with prior research identified in this literature review, but also offers some unique insights.

For starters, "generic" mass public shootings do not tend to occur in major metropolitan or urban areas, yet Black Swan attacks occur in above average population densities and show a relationship to sudden changes in population. One of the explanations for the relationship between economic inequality from Kwon and Cabrera (2019) is "income inequality fosters an environment of anger and resentment that ultimately leads to violence." Similar logic could apply to the relationship between changes to population density and Black Swan Shootings. Perpetrators of general mass violence are overwhelmingly white males, yet Black Swan attacks occur in more heterogeneous counties than average across the U.S., and Black Swan shooters live in communities with less non-Hispanic Caucasians than average. Black Swan attack and residence counties have above averages increases in female population during the year of a shooting. Coupled with the measures for race, and the anecdotal information about incel, misogynistic, racist beliefs, and toxic masculinity (Follman, 2019) of many mass shooters, there is an inference to be made about relative changes to a community that put the offender in an irrationally heightened state. Both measures of race and gender tap into the notion of White males feeling uncomfortable, victimized, and disenfranchised by community deviations from expected norms. These differences create some of the conditions conducive for mass violence. Similarly, while violence crime does not demonstrate relationships, LEOKA does. In this study, LEOKA serves as a proxy for tension and hostility towards law enforcement. These findings indicate a potentially eroded relationship between law enforcement and the public in these communities. It has become exceedingly well documented that particular firearm legislation limiting access to at-risk individuals and requiring common-sense background checks and waiting periods has positive effects on firearm violence. Yet, this study finds Black Swan Shootings occur in areas with more firearm laws in place at the time of the attack. This is largely driven by California, where there are some of the highest counts of firearm laws in the U.S. and the most Black Swan Shootings. Future studies could explore this further and identify whether this is a spurious relationship or an additional environmental agitant. Ultimately, in a larger context, these social factor relationships to Black Swan Shooting activity highlight the relative unrest or tension in a community.

Third, this study has attempted to develop a threat assessment model for identifying high-risk counties each year. Knowing the social factors have these distinct relationships, can they be modeled to effectively flag areas at highest risk for an attack? The answer here also appears to be yes. For Black Swan attack counties, the logistic regression model identifies approximately 21 counties each year at highest risk for a Black Swan Shooting. This is less than 1% of the entire country. And while the attack probability for a high-risk county is still low (less than 4%, on average), since 2006 the model has been analytically viable ten different years. This includes eight bullseyes, where the county and year of a Black Swan attack were exact matches. These bullseyes represent 30% of Black Swan Shootings since 2008. Briefly, here are some of the details of the exact matches the model retroactively identified:

- 12/24/2008, Covina, CA: single family home, 11 total victims
- 11/5/2009: Fort Hood, TX: Fort Hood Military Base, 43 total victims
- 10/12/2011, Seal Beach, CA: Salon Meritage Hair Salon, 9 total victims
- 4/3/2014: Fort Hood, TX: Fort Hood Military Base, 15 total victims
- 12/2/2015, San Bernardino, CA: Inland Regional Center, 35 total victims
- 7/7/2016, Dallas, TX: Black Lives Matter March, 16 total victims
- 9/10/2017: Plano, TX: single family home, 10 total victims
- 10/1/2017: Las Vegas, NV: Route 91 Harvest Music Festival, 500 total victims

It is worth noting that five of these attacks (both Fort Hood attacks, San Bernardino, Black Lives Matter, and Route 91 in Las Vegas) are some of the most well-known, highly publicized mass shootings in American history. These results highlight the benefit of the approach employed in this study: these attacks are not related to each other. They are not part of a larger network, organization, or serial offender; rather, these are isolated attacks committed by different perpetrators for different reasons. However, as this study contends, it is not the motivation or traits of the offender that are predictable here; rather, it is the location-based features that made these communities vulnerable to attacks. Among these attack bullseyes, on average the communities have consistently high populations, high population densities, high LEOKA rates, low percentages of Caucasian population, high percentages of females, low drug and alcohol overdose rates, low percentage of suicides committed by firearm, and high counts of firearm laws. Measures for crime rate, poverty, unemployment, and suicide vary across these community bullseyes.

For Black Swan residence counties, the results are similarly compelling. Annually, there are on average 13 communities at highest risk for potential Black Swan residences. While there are less bullseyes from the model (six versus eight in the attack model), the relative spatial accuracy is much closer for residences. The calculated spatial distance is half as far as the attack model (Table 34), suggesting that the location-based factors in this model are more consistent. The attack model appears to be high risk, high reward, getting the county and year exactly right multiple times, right but being more prone to big misses. The residence model is lower risk-lower reward, with the ability to get to the correct general area, but not the exact location, consistently. Some of this can be attributed to using the spatial and temporal contagion measures in the attack model. Those effects make the model better when attacks are consistent. However, when annual fluctuations change (and drop to zero attacks in a given year), those effects appear to negatively impact the results.

Thus, it does seem possible to identify high-risk areas for the worst of the worst mass shootings. The patterns are not inherently obvious nor offender-centric, but nonetheless exist. The logistic regression models have at a minimum created a small enough hotlist of communities (of less than 1% of the country) and has been exactly right and had enough near misses for attack and residence locations, to be deemed a success.

And while the models are certainly far from perfect, they are encouraging. This analysis suggests that subconscious decision-making is real, that the *where* of mass violence matters, and that certain areas of the country have an environment more conducive for such attacks. Like other crime prevention methods, this study indicates that focusing on the where of mass violence has potential implications for threat mitigation and prevention, and that the threat assessment perspective has merit for mass violence.

Limitations

As discussed, this study's greatest limitation is its low number of events. It is prone to high volatility, large variations, and unexplainable relationships. While the method for selecting Black Swan Shootings was thorough, comprehensive, and defensible, it may have been too restrictive. A sensitivity analysis of mass shooting events, as shown in Table 36, highlights the possibilities. Using the same timeline of events as the current study (1998), dropping the threshold to as low as seven total casualties would capture approximately 301 events. Other casualty counts, such as five killed (204 events), six killed (98 events), and nine wounded (71 events) capture significantly more shootings. These thresholds were not used in this study for the reasons outlined in the methodology; despite having greater counts than “ordinary” mass shootings, these were not deemed rare enough to Black Swans. However, future studies that adjust casualty thresholds have the opportunity for an expanded analysis and different findings. Additional research is necessary to identify the extent of these findings to a less restrictive mass shooting dataset.

Table 36 Black Swan Shooting sensitivity analysis

		Thresholds		
	Event Count	**Killed**	**Wounded**	**Total**
This study	44	8	13	15
Since 1998	301			7
	204	5		
	187			8
	131			9
	102			10
	98	6		
	71		9	
	58	7		
	47		11	
	45	8		
	42		12	
	39		13	

***all event counts are approximates, without further data cleaning and merging**

Some of this is exacerbated by a second limitation, incomplete social factors data. In this study, some data sources are used out of necessity, despite their errors, gaps, and omissions. Given the timing of this study, violent crime and LEOKA data for 2019 are not yet available from the FBI. The 2020 Census has recently begun, which will soon generate new socioeconomic data. CDC data for suicides and drug and alcohol overdoses for 2019 are not yet available. But even for the available CDC data, there appear to be gaps in coverage, with large amounts of missing and incomplete county counts. This appears to be not widespread, but additional depth and breadth of data across more factors and more years would potentially resolve some of these issues. Data on firearm legislation is the most clear and consistent, due to the study design. By only capturing the number of firearm laws, data covers the entire study period. However, qualitative judgements such as report cards are not studied. By examining the aggregate counts, this study misses the opportunity to analyze the types of laws showing the strongest relationships to Black Swan Shootings. Given the complexities with each of the social factor datasets, the amount of variation across the statistical testing is sometimes low. Some of these relationships may be correlation, not causation.

A third limitation is the study design. An alternative approach could involve a one-year lag for event-driven analytic layers. This would take social factors such as crime, law enforcement violence, and substance abuse, and use the year *prior* for analyzing a given year. It would theoretically capture a full year of activity leading into the event. It would also allow for a prediction of more recent activity; using 2018 data, predictions for 2019 could be made. This was not done in the current study for several reasons. First, this study was concerned that a mismatch of lag and current year social factor analysis would be difficult to interpret and extrapolate findings. Second, the substance abuse data from CDC is inconsistent and limited. Querying data a year earlier was expected to cause further data quality and model performance issues. A future study with different social factor data sources could potentially involve a more systematic approach to addressing lagged variables, to include multi-year effects. That complexity was outside the scope of this research.

Further, a larger event dataset could have changed some of the modeling techniques. Predictive modeling and machine learning principles typically involve building training and test data from the dataset. The training data is used for creating the model, the test data is used to assess model performance, and then predictions are made. This study did not use this methodology, due to the small event dataset. Future studies with lower thresholds and more events may benefit from such techniques.

Further, the logistic regression models used in this study were built on some variables that were not statistically significant. Despite the rigorous testing across variations of the fields and multiple statistical measures that demonstrated statistically distinct and significant relationships, when used in the regression equations (Tables 30 and 31) they returned larger p-values. This does not negatively influence the results, but it is worth considering adjusting the modeling techniques for future research.

One final consideration here is that the mass shooting events in this study are anomalies. Even within the context of mass violence, Black Swan Shootings are an even smaller subset of violence. And by their anomalous nature, they have greater potential to fall outside of conventional wisdom and preconceived notions. This is apparent in some of the relationships to different social factors. As the literature review demonstrated, strict laws regulating firearms can prevent firearm violence, including mass shootings. Among Black Swan Shootings, high-risk communities have significantly higher amounts of firearm laws. Spatial contagion effects provide a potential explanation. Firearm laws are a state measure. Communities adjacent to previous Black Swan Shootings are often in the same state. State and federal legislative responses, albeit slow, do occur over time. Thus, current high-risk areas may have a combination of proximity to prior attacks, as well as higher numbers of firearm laws. Or, states like California are truly an anomalous mix of firearm laws and violence. Appreciating the nuance of addressing Black Swan Shootings requires special focus and attention, which leads to potential implications of this research.

Implications

This study demonstrates the potential location-based predictability of rare mass violence events. It rebukes the notion that such extreme mass shootings are truly "unpredictable," but occur in locations that share a small set of similar, identifiable features. These findings can serve as actionable intelligence for planning, interdicting, and responding to such events. For any prediction to be tactical relevant, it must be precise enough to be utilized. The average precision for the high-risk attack and residence Black Swan models was 99%. The attack modeled identifies 21 high-risk counties each year; the residence model finds 13. These results are precise enough to be incorporated into tangible action plans.

But what could an action plan using this study look like? In a multi-pronged approach that involves federal, state, and local government agencies across law enforcement, education, social work, and medicine, this model is the initial layer. If coordinated action to prevent mass violence is viewed as a funnel, the threat assessment in this study is at the top. Applying this assessment to a given year, it removes 99% of the "noise" across the country. Instead of uncoordinated, inconsistent local responses, this model can spatially filter and allocate centralized resources to the communities at highest risk for extreme mass violence. Thus, step one is identifying those areas.

Step two is to fuse this model with individual and behavioral threat assessments at the local level. Once an area is prioritized, government agencies can begin targeted intervention and investigation strategies. For example, instead of the daunting and largely impossible task for federal law enforcement to scan the entire country to identify mass violence threats in a top-down approach, they could instead start in a handful of prioritized communities. This is more practical and allows for traditional law enforcement work (source development, tips and leads, investigative research) and newer methods (cyber and electronic tracking, for example) to be applied, in a bottom-up approach. For law enforcement, this would enable collaboration and coordination across local police, state agencies, and the FBI. It would also enable better communication for school threats, suspicious health behaviors, and increased community involvement. Federal resources diverted to high-risk communities would rely on the organic knowledge and access to information from local resources.

Such a plan would not ignore threats from lower risk communities; rather, it would initially prioritize high-risk communities, in the absence of specific/credible threat information. All levels of government have limited resources, and no single discipline is designed to identify, resolve, interdict, and investigate threats on a nationwide scale. This model is a deductive analytic, that filters to a significantly smaller subset. Working from macro to micro, this macro-geographic analytic is the beginning phase, that can be followed by additional analysis, research, and investigation into micro-geographic and eventually identification of individuals of interest. While this study demonstrates the value of location-based threat assessment, it does not eliminate the need for offender/behavioral threat assessments. This study seeks to maximize the value of that work in targeted locations. As Silver (2020) described, threat assessments are not profiling or offender-oriented at all; rather, it is event-driven and target-focused. Threat assessment is built for statistically rare events. This study seeks to be a part of the process, not the sole answer to the problem.

Law enforcement collaboration for narcotics has a similar macro-micro collaborative environment. The Department of Justice Drug Enforcement Agency has established regional task forces, known as High Intensity Drug Trafficking Areas (HITDA) to foster data sharing and combined investigations and operations to address the areas at highest risk for narcotics problems. Thus, study advocates for a similar model: create local task forces for high-risk mass violence communities, where federal and local combine to identify and disrupt threats.

However, this model alone is worthless without proper integration into other practices, disciplines, intelligence, and related materials. This work does not stand alone, as much as it serves to supplement other work in this field. While the predictive results are promising, they are not perfect; multiple attacks go initially undetected in this model. Other disciplines and measures used in conjunction with this process may offer additional insights into those relative unknowns.

Further, the takeaway here should not be that more laws equal more mass shootings. Yes, there is a correlation in this study; but that correlation exists among the 44 most severe attacks in the past 21 years. It does not account for the hundreds of other mass violence attacks that occurred. These findings should not be extrapolated to all forms of mass violence, where more comprehensive studies have found significant reductions in violence from increased firearm legislation. As discussed earlier, a different threshold for extreme activity is expected to demonstrate different relationships.

Building off of Nagin et al. (2020), this study has the potential to improve formal tracking mechanisms for mass violence; improving threat detection systems to better detect and respond to 'leakage'; and obviously expand and improve use of threat assessment teams. And similar to the work of Berk and Sorenson (2020), this study hopes to be able to be part of the dialogue for developing better tools to assess potential threats. Thus, updated data should be obtained to run this model again, to begin the process of identifying high-risk communities for future years.

Directions for future research

What gaps in mass violence research does this study fill? This study takes prior research on mass violence from studies such as Markowiak et al. (2018), Webster et al. (2016), Reeping et al. (2019), Kwon and Cabrera (2019), and Follman (2019) and builds community and location-based correlations for mass shootings. This study confirms some of the known relationships to mass violence, in particular socioeconomic measures such as population density. This study also sheds light on new relationships, such as LEOKA, race, gender, and state firearm laws. In both cases, this study frames correlations and relationships as environmental factors that make a community conducive for an attack to occur. Further, this study shows the utility of the threat assessment model for a topic that has been difficult to anticipate and predict. This study demonstrates how threat assessment and degrees of risk probability, analyzed at the community level, may be more effective than offender-centric targeting for modeling, understanding, and interdicting mass violence.

What gaps in mass violence research still exist? There are many directions for future research. First and foremost, there is the definition of a Black Swan Shooting. As discussed in the limitations section, variations of the attack definition, coupled with different victim thresholds and temporal start points, has the potential for countless other studies. A larger opportunity exists for taking these methods and increasing the scale and parameters of the events. There is potential for lower thresholds for victims, and different attack types. Black Swans do not necessarily need to be limited to shootings; future studies may address other types of violence.

Second, there are additional and alternative methods for measuring contagion. More analysis of the cluster and spree dynamics identified here is warranted. If these contagion effects are real, those patterns should be monitored for changes over time. Alternative methods include silhouette analysis, gap statistics, hierarchical clustering, and ordering points to identify the clustering structure (OPTICS) analysis. There is the potential to better identify the similarities and differences in attack behaviors between attacks in the same hotspots. Future studies can measure hotspot changes over time and see if the Midwest and non-coastal hotspots truly remain inactive. Primary sprees were relatively useless in this study; but since two of the four have occurred since late 2017, they are worth monitoring and re-assessing with new attack data in a future study. Further studies should complement the macro-geographic focus of this study with an analysis of the micro-geographic factors of such attacks, to better understand how specific attack location types are influenced by surrounding features. This should involve a Risk Terrain Modeling perspective.

Third, there are different social factors to measure. There are opportunities for further correlating Black Swan Shootings to other location-based, social, political, and environmental indicators. New socioeconomic measures to study include the traditional metrics, such as education level, divorce rate, and head of household rate. This study did not find relationships with Black Swan Shootings and poverty and unemployment. However, poverty and unemployment are proxies for housing and food insecurity. Poverty and unemployment may be too generic, and better data on community housing costs, food costs, wealth disparity, and similar measures could identify better relationships. Refined socioeconomic measures include a different definition for race. Identifying non-Hispanic white alone individuals is one way to measure diversity, but there are countless others. While this does not inherently measure heterogeneity or diversity in an area, it does measure ‘whiteness’, and it can be reasonably inferred that the remaining percentage of people in a county are indicative of the amount of diversity for that community. A racial heterogeneity measure may yield different results.

New violence measures could include traditional statistics on hate crime, law enforcement-perpetrated violence (which is growing in data collection), and other terrorist attacks. However, violence may be best captured through non-traditional data sources. Domestic violence restraining orders, as a count, rate of the population, and percentage of all restraining orders in a community could be insightful. Extrapolating factors from criminal histories for mass violence perpetrators in general, and specifically Black Swan Shooters, could highlight additional characteristics to study. Further, there is social media. A future study can and should intelligently leverage Internet-based content. This could be in the form of modeling sentiment of prior perpetrators to identify new threatening individuals, based on social media or blog posts. It could also be monitoring chat rooms for hate groups and the incel subculture, and collecting device metadata (IP, MAC addresses) for visitors. Future studies may be inclined to pursue this relationship further and examine specific offenders and a desire to commit any racially motivated attacks.

For firearm preferences, aside from higher quality suicide data, directions for different or enhanced data include coding firearm laws differently, capturing voter preferences, and political party affiliations. Data on firearm law report cards by state from the Giffords Law Center to Prevent Gun Violence exists for 2013-2018. However, given the abbreviated time range (six years), and the vastly different measurement scale (A through F plus/minus academic ranks), the more straightforward count measure was favored in this study. In the future, coding firearm laws would involve not counting the number of laws, but contextualizing the laws by type, restrictiveness, uniqueness, and responsiveness to mass violence. This could help us understand firearm laws as a sort of temporal contagion, identifying the states that respond through legislation to prior mass violence attacks. Voter preferences and political party distributions are additional proxies for firearm preferences. Future studies could include Congressional voting records on firearm legislation by district, primary election results for candidates grouped by their firearm politics, and community surveys on weapons control. This study scratches the surface on community risk factors and hopes to contribute to an ever-growing body of literature on this topic but recognizes there are many other datasets to analyze on this topic.

Fourth, analytically there are different modeling techniques to test. Aside from the sensitivity analysis and subsequent machine learning methods discussed earlier, future studies should leverage enhanced variable selection methods, such as least absolute shrinkage and selection operator (LASSO). Such techniques eliminate the need for complex methodologies and generic 0.05 cut-offs. This study did not use LASSO because the variable selection was a quantitative and qualitative review of multiple measures that led to a reasonable set of factors to model. There are different modeling techniques besides logistic regression, including random forest learning models and machine-learning techniques. And even without changing event parameters and social factors analyzed here, there is further analysis to be done on these events. This includes detailed location analysis; correlations across casualty counts, offender count, race, and age; weapon use and weapon type correlations and analysis, to include assault weapons and high capacity magazines; victim-offender relationships' influence on site selection; and the relationship of mental health indicators to target selection, casualty counts, and community-level measures for mental health.

There are clearly many possible directions for future research on this topic. The intent here is to further the dialogue about mass violence, identify some location-based patterns that offer predictive potential, demonstrate what a threat assessment model that contribute to the dialogue, and offer tangible potential solutions to a disturbingly increasing problem.

REFERENCES

Addington, L.A. (2004). The effect of NIBRS reporting on item missing data in murder cases. *Homicide Studies, 8:* 193-213, doi: 10.1177/1088767904265360.

Aitken, L.; Oosthuizen, P.; Emsley, R.; Seedat, S. (2008). Mass murderers: Implications for mental health professionals. *International Journal of Psychiatry in Medicine, 38*(3): 261-269.

Alathari, L.; Drysdale, D.; Blair, A.; McGarry, J.; Camilletti, C.; Snook, A.; & Driscoll, S. (2018). Enhancing school safety using a threat assessment model: An operational guide for preventing targeted school violence. *U.S. Department of Homeland Security, United States Secret Service, National Threat Assessment Center*, July 2018. Available at https://www.secretservice.gov/data/protection/ntac/USSS_NTAC_Enhancing_School_Safety_Guide_7.11.18.pdf.

Allely, C.S.; Minnis, H.; Thompson, L.; Wilson, P.; Gillberg, C. (2014). Neurodevelopmental and psychosocial risk factors in serial killers and mass murderers. *Aggression and Violent Behavior, 19(3)*: 288-301, doi:10.1016/j.avb.2014.04.004.

Altheimer, I.; Boswell, M. (2012). Reassessing the association between gun availability and homicide at the cross-national level. *American Journal of Criminal Justice*, 37(4): 682-704, doi 10.1007/s12103-011-9147-x.

Andresen, M.A.; Malleson, N. (2013). Crime seasonality and its variations across space. *Applied Geography, 43*: 25–35.

Aneja, A.; Donohue, J.J.; Zhang, A. (2014). The impact of right to carry laws and the NRC report: The latest lessons for the empirical evaluation of law and policy. Stanford Law and Economics Online Working Paper No. 461, doi: 10.2139/ssrn.2443681.

Anestis, M.D.; Anestis, J.C. (2015). Suicide rates and state laws regulating access and exposure to handguns. *American Journal of Public Health, 105(10)*: 2049-2058, doi:10.2105/AJPH.2015.302753.

Anglemyer, A.; Horvath, T.; Rutherford, G. (2014). The accessibility of firearms and risk for suicide and homicide victimization among household members: A systematic review and meta-analysis. *Annals of Internal Medicine*, 160(2):101-110, doi: 10.7326/M13-1301.

Anisin, A. (2019). Mass shootings and their asymmetric effect on societal armament. Crime, Law, and Social Change, doi: https://doi-org.mutex.gmu.edu/10.1007/s10611-019-09832-x.

Anselin, L.; Cohen, J.; Cook, D.; Gorr, W.; Tita, G. (2000). Spatial analyses of crime. In *Measurement and analysis of crime and justice: Criminal justice*, Edited by: Duffee, D. Volume 4: 213–262. Washington DC: National Institute of Justice.

Ardohain, C.M. (2016). IED pattern recognition using sinusoidal models. *Calhoun Institutional Archive of the Naval Postgraduate School*. Available at https://calhoun.nps.edu/bitstream/handle/10945/49454/16Jun_Ardohain_Christopher.pdf.

Ayres, I.; Donohue, J.J. (2003). Shooting down the more guns, less crime hypothesis. *Faculty Scholarship Series,* paper 1241. Available at http://digitalcommons.law.yale.edu/fss_papers/1241.

Azrael, D.; Cook, P.J.; Miller, M. (2004). State and local prevalence of firearms ownership: measurement, structure, and trends. *Journal of Quantitative Criminology, 20(1):* 43–62.

Baker, J.; McPhedran, S. (2007). Gun laws and sudden death: Did the Australian firearms legislation of 1996 make a difference? *British Journal of Criminology, 47(3)*: 455-69.

Berk, R.A.; Sorenson, S.B. (2019). Algorithmic approach to forecasting rare violent events: An illustration based in intimate partner violence perpetration. *Criminology and Public Policy, 19*(1): doi: https://doi.org/10.1111/1745-9133.12476.

Bernard, T.J; Snipes, J.B.; Gerould, A.L. (2015). Neighborhoods and crime. *Vold's Theoretical Criminology*. Oxford University Press.

Bisakha, S.; Panjamapirom, A. (2012). State background checks for gun purchase and firearm deaths: an explanatory study. *Preventive Medicine*, 55(4): 346–350.

Bjelopera, J.P.; Bagalman, E.S.; Caldwell, W.; Finklea, K.M.; McCallion, G. (2013). Public mass shootings in the United States: Selected implications for federal public health and safety policy. *Congressional Research Service*. Available at https://fas.org/sgp/crs/misc/R43004.pdf.

Blair, J.P.; Schweit, K.W. (2014). A study of active shooter incidents, 2000–2013. *Texas State University* and *Federal Bureau of Investigation, U.S. Department of Justice*. Available at https://www.fbi.gov/news/stories/2014/september/fbi-releases-study-on-active-shooter-incidents/pdfs/a-study-of-active-shooter-incidents-in-the-u.s.-between-2000-and-2013.

Blau, J.; Blau, P. (1982). The cost of inequality: Metropolitan structure and violent crime. *American Sociological Review*, 47(1): 114–129.

Boba Santos, R. (2013). *Crime analysis with crime mapping*. SAGE publications.

Boor, M.; Bair, J.H. (1990). Suicide rates, handgun control laws, and sociodemographic variables. *Psychol Rep*, 66: 923-930.

Bowers, T.G.; Holmes, E.S.; Rhom, A. (2010). The nature of mass murder and autogenic massacre. *Journal of Police Crime Psychology 25*, doi: 10.1007/s11896-009-9059-6.

Brady Campaign To Prevent Gun Violence. (2019). Mass shooting data. Available at: https://www.bradyunited.org/.

Braga, A.; Weisburd, D. (2012). The effects of focused deterrence strategies on crime: A systematic review and meta-analysis of the empirical evidence. *Journal of Research in Crime and Delinquency*, 49(3): 323-358, doi: 10.1177/0022427811419368.

Brantingham, P.L.; Brantingham, P.J. (1993). Environment, routine, and situation: Toward a pattern theory of crime. In *Routine Activity and Rational Choice: Advances in Criminological Theory*, R.V. Clarke & M. Felson (Eds.), 5: 259-294.

Briggs, J.T.; Tabarrok, A. (2013). Firearms and suicides in U.S. states. *George Mason University*. Available at http://mason.gmu.edu/~atabarro/BriggsTabarrokFirearmsSuicide.pdf.

Buchanan, L.; Keller, J.; Oppel, Jr., R.A.; Victor, D. (2015). How they got their guns. *New York Times*. Available at http://www.nytimes.com/interactive/2015/10/03/us/how-mass-shooters-got-their-guns.html.

Burdick-Will, J. (2018). Neighborhood violence, peer effects, and academic achievement in Chicago. *Sociology of Education*, 91(3): 205-223, doi: 10.1177/0038040718779063.

Busch, K.A.; Cavanaugh, J.L., Jr. (1986). The study of multiple murder: Preliminary examination of the interface between epistemology and methodology. *Journal of Interpersonal Violence, 1(1)*: 5-23.

Cannon, A. (2016). Mayhem multiplied: Mass shooters and large capacity magazines. *Citizens Crime Commission of New York City*. Available at http://www.nycrimecommission.org/pdfs/CCC-MayhemMultiplied-June2016.pdf.

Cantor, C.H.; Mullen, P.; Alpers, P. (2000). Mass homicide: The civil massacre. *Journal of the American Academy of Psychiatry and the Law, 28(1)*: 55-63.

Cantor, C.H.; Sheehan, P.; Alpers, P.; Mullen, P. (1999). Media and mass homicides. Archives of Suicide Research, 5(4): 283-290,doi: 10.1023/A:1009637817185.

Cantor, C.H.; Slater, P.J. (1995). The impact of firearm control legislation on suicide in Queensland: Preliminary findings. *Medical Journal of Australia, 162*: 583-585.

Caplan, J.M.; Kennedy, L.W. (2016). *Risk terrain modeling: Crime prediction and risk reduction*. University of California Press.

Carcach, C.; Mouzos, J.; Grabosky, P. (2002). The mass murder as quasi-experiment: The impact of the 1996 Port Arthur massacre. *Homicide Studies, 6(2)*: 109-127.

Chaiken, J.; Chaiken, M.; Rhodes, W. (1994). Predicting violent behavior and classifying violent offenders. In *Understanding and preventing violence*, A. Reiss & J. Roth (Eds.), 4: 217-295. Washington, DC: National Academy Press, available at https://www.nap.edu/read/4422/chapter/5.

Chainey, S.; Tompson, L.; Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal, 21*: 4–28.

Chapman, S.; Alpers, P.; Agho, K.; Jones, M. (2006). Australia's 1996 gun law reforms: Faster falls in firearm deaths, firearm suicides, and a decade without mass shootings. *Injury Prevention, 12*: 365-372.

Chapman, S.; Alpers, P.; Jones, M. (2016). Association between gun law reforms and intentional firearm deaths in Australia, 1979-2013. *Journal of the American Medical Association, 316(3)*, doi: 10.1001/jama.2016.8752.

Citizens Crime Commission of New York City. (2012). Mass shooting incidents in America (1982-2012), available at: http://www.nycrimecommission.org/mass-shooting-incidents-america.php.

CNN (2020). Deadliest Mass Shootings in U.S. History Fast Facts. Available at http://www.cnn.com/2013/09/16/us/20-deadliest-mass-shootings-in-u-s-history-fast-facts/.

Coe. (2002). It's the effect size, stupid: What effect size is and why it is important. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England. Available at https://www.leeds.ac.uk/educol/documents/00002182.htm.

Cohen, A.P.; Azrael, D.; Miller, M. (2014). Rate of mass shootings has tripled since 2011, Harvard research shows. *Mother Jones*, available at http://www.motherjones.com/politics/2014/10/mass-shootings-increasing-harvard-research.

Cohn, D.; Taylor, P.; Hugo Lopez, M.; Gallagher, C.A.; Parker, K.; Mass, K.T. (2013). Crime News. *Pew Research Center*, available at: http://www.pewsocialtrends.org/2013/05/07/chapter-5-context/.

Cook, P.J.; Ludwig, J. (1997). Guns in America: National survey on private ownership and use of firearms. National Institute of Justice. Available at https://www.ncjrs.gov/pdffiles/165476.pdf.

Covucci, D. (2016). The short history of every time Congress tried, and failed, to prevent a mass shooting. *Huffington Post.* Available at: http://www.huffingtonpost.com/entry/congress-guns-mass-shootings_us_5850547be4b04c8e2bb258a3.

Crifasi, C.K.; Merrill-Francis, M.; McCourt, A.; Vernick, J.S.; Wintemute, G.J.; & Webster, D.W. (2018). Association between firearm laws and homicide in urban counties. Journal of Urban Health, 95(3): 383-390.

D'Anna, M.E. (2016). Social media perceptions of firearms and U.S. mass shootings: A comparative analysis. George Mason University graduate research paper.

Dahmen, N.S.; Abdenour, J.; McIntyre, K.; Noga Styron, K.E. (2017). Covering mass shootings. *Journalism Practice*, doi: 10.1080/17512786.2017.1326832.

Declercq, F.; Audenaert, K. (2011). A case of mass murder: Personality disorder, psychopathology and violence mode. *Aggression and Violent Behavior, 16*: 135-143, doi: 135–143.10.1016/j.avb.2011.02.001.

Dezhbakhsh, H.; Rubin, P.H. (1998). Lives saved or lives lost? The effects of concealed handgun laws on crime. *American Economic Review, 88*: 468-474.

Dietz, P.E. (1986). Mass, serial, and sensational homicides. *Bulletin of the New York Academy of Medicine, 62:* 477-490.

Dillon, L. (2013). Mass shootings in the United States: An exploratory study of the trends from 1982-2012. Thesis, George Mason University.

Donohue, J.J.; Aneja, A.; Weber, K.D. (2018). Right-to-carry laws and violent crime: A comprehensive assessment using panel data, the LASSO, and a state-level synthetic controls analysis. *The National Bureau of Economic Research*, doi: 10.3386/w23510.

Dowden, C. (2005). Research on multiple murder: Where are we in the state of the art? *Journal of Police and Criminal Psychology, 20 (2):* 8-19.

Duggan, M. (2001). More guns, more crime. *The Journal of Political Economy, 109(5)*: 1086-1114.

Duong, T. (2007). Ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. *Journal of Statistical Software, 21(7),* available at: https://pdfs.semanticscholar.org/473a/9239ade9085ad5fd0206fcbb22d17f5571c7.pdf.

Dutton, D.G.; White, K.R.; Fogarty, D. (2013). Paranoid thinking in mass shooters. Aggression and Violent Behavior, 18: 548-553, doi: 10.1016/j.avb.2013.07.012.

Duwe, G. (2000). Body-count journalism: The presentation of mass murder in the news media. *Homicide Studies, 4*(4)*:* 364-399.

Duwe, G. (2004). The patterns and prevalence of mass murder in twentieth-century America. *Justice Quarterly 21(4):* 729–761. doi: 10.1080/07418820400095971.

Duwe, G. (2005). A circle of distortion: The social construction of mass murder in the United States. *Western Criminology Review 6*(1): 59–78.

Duwe, G. (2007). *Mass murder in the United States: A history*. McFarland Press.

Duwe, G. (2019). Patterns and prevalence of lethal mass violence. *Criminology and Public Policy, 19(1)*, doi: https://doi.org/10.1111/1745-9133.12478.

Duwe, G.; Kovandzic, T.; Moody, C.E. (2002). The impact of right-to-carry concealed firearm laws on mass public shootings. *Homicide Studies, 6(4)*: 271-296.

Duxbury, S.W.; Frizzell, L.C.; Lindsay, S.L. (2018). Mental Illness, the Media, and the Moral Politics of Mass Violence: The Role of Race in Mass Shootings Coverage. Journal of Research in Crime and Delinquency, 55(6), doi: https://doi.org/10.1177%2F0022427818787225.

Elbogen, E.B.; Johnson, S.C. (2009). The intricate link between violence and mental disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry, 66(2)*: 152-161, doi: 10.1001/archgenpsychiatry.2008.537.

Everytown. (2020). Mass shootings in the United States: 2009-2019. Available at http://everytownresearch.org/reports/mass-shootings-analysis/.

Farr, K. (2017). Adolescent rampage school shootings: Responses to failing masculinity performances by already-troubled boys. Gender Issues, October 2017, doi: 10.1007/s12147-017-9203-z.

Federal Bureau of Investigation [F.B.I.]. (2018). Active Shooter Incidents in the United States in 2016 and 2017. Available at https://www.fbi.gov/file-repository/active-shooter-incidents-us-2016-2017.pdf/view.

Federal Bureau of Investigation. Annual reports: Crime in the United States, 1999-2018. U.S. Department of Justice. http://www.fbi.gov/about-us/cjis/ucr/ucr-publications#Crime.

Federal Bureau of Investigation [F.B.I.]. (2020). National Incident-Based Reporting System (NIBRS). Available at: https://www.fbi.gov/services/cjis/ucr/nibrs.

Federal Bureau of Investigation [F.B.I.]. (2019). NICS Firearm Checks: Month/Year. Available at https://www.fbi.gov/file-repository/nics_firearm_checks_-_month_year.pdf/view.

Fegadel, A.R.; Heide, K.M. (2016). NIBRS and SHR: A comparison of two national homicide databases with respect to parricide. *Victims and Offenders*, 1-22, doi: 10.1080/15564886.2016.1246392.

Ferguson, C.J.; Coulson, M.; Barnett, J. (2011). Psychological profiles of school shooters: Positive directions and one big wrong turn. *Journal of Police Crisis Negotiations, 11*: 1-17, doi: 10.1080/15332586.2011.581523.

Fitzgerald, M. (2015). Autism and school shootings — Overlap of autism (Asperger's syndrome) and general psychopathy. Available at http://cdn.intechopen.com/pdfs-wm/47518.pdf, doi: 10.5772/58882.

Flannery, D.J.; Modzeleski, W.; Kretschmar, J.M. (2013). *Current Psychiatry Reports, 15(331)*: 1-7, doi: 10.1007/s11920-012-0331-6.

Fleegler, E.W.; Lee, L.K.; Monuteaux, M.C.; Hemenway, D.; Mannix, R. (2013). Firearm legislation and firearm-related fatalities in the United States. *Journal of the American Medical Association, 173*: 732–740. doi: 10.1001/jamainternmed.2013.9957. Pmid:23467753.

Florida, R. (2011). The geography of gun deaths. *The Atlantic.* Available at https://www.theatlantic.com/national/archive/2011/01/the-geography-of-gun-deaths/69354/.

Follman, M. (2015). Inside the race to stop the next mass shooter. *Mother Jones.* Available at: https://www.motherjones.com/politics/2015/10/mass-shootings-threat-assessment-shooter-fbi-columbine/.

Follman, M. (2019). Armed and misogynist: How toxic masculinity fuels mass shootings. *Mother Jones.* Available at: https://www.motherjones.com/crime-justice/2019/06/domestic-violence-misogyny-incels-mass-shootings/.

Follman, M.; Aronsen, G.; Pan, D. (2020). US mass shootings, 1982-2020: Data from Mother Jones' investigation, available at: https://www.motherjones.com/politics/2012/12/mass-shootings-mother-jones-full-data/.

Follman, M.; Pan, D.; Aronsen, G. (2018). A guide to mass shootings in America. *Mother Jones*. Available at http://www.motherjones.com/politics/2012/07/mass-shootings-map/.

Ford, C. (2017). The Wilcoxon Rank Sum Test. *Research Data Services and Sciences, University of Virginia Library*, available at: https://data.library.virginia.edu/the-wilcoxon-rank-sum-test/.

Fox, J.A. (2013). Mass shootings not trending. Available at http://archive.boston.com/community/blogs/crime_punishment/2013/01/mass_shootings_not_trending.html.

Fox, J.A.; DeLateur, M.J. (2014). Mass shootings in America: Moving beyond Newtown. *Homicide Studies, 18(1)*: 125-145, doi: 10.1177/1088767913510297.

Fox, J. A.; Levin, J. (1994). Firing back: The growing threat of workplace homicide. *Annals of the American Academy of Political and Social, 536*: 16-30.

Fox, J.A.; Levin, J. (1998). Multiple homicide: Pattern of serial and mass murder. *Crime and Justice, 23:* 407-455.

Fox, J.A.; Levin, J. (2003). Mass murder: An analysis of extreme violence. *Journal of Applied Psychoanalytic Studies, 5:* 47-64.

Fox, J.A.; Levin, J. (2015). Mass confusion surrounding mass murder. *The Criminologist, 40*: 8–11.

Fridel, E.E. (2019). The Contextual Correlates of School Shootings. Justice Quarterly, doi: https://doi.org/10.1080/07418825.2019.1666907.

Geisel, M.S.; Roll, R.; Wettick, R.S. (1969). The effectiveness of state and local regulations of handguns. *Duke University Law Journal, 4*: 647–676.

Gill, P.; Silver, J.; Horgan, J.; Corner, E. (2016). Shooting alone: The pre-attack experiences and behaviors of U.S. solo mass murderers. *Journal of Forensic Sciences, 62(3)*: 710-714, doi: 10.1111/1556-4029.13330.

Gius, M. (2015). The impact of state and federal assault weapons bans on public mass shootings. *Applied Economics Letters, 22(4)*: 281–284.

Gottfredson, S.; Gottfredson, D. (1988). Violence prediction methods: Statistical and clinical strategies. *Violence and Victims 3(4),* 303-324.

Goode, E.; Ben-Yehuda, N. Moral panics: Culture, politics, and social construction. *Annu. Rev. Sociol., 20*: 149-171.

Gresswell, D.M.; Hollin, C.R. (1994). Multiple murder: A review. *The British Journal of Criminology, 34*(1): 1-14.

Groff, E. R.; La Vigne, N. G. (2001). Mapping an opportunity surface of residential burglary. *Journal of Research Crime and Delinquency, 38*: 257–278.

Gun Violence Archive (2019). Available at http://www.gunviolencearchive.org/.

Guo, J. (2015). Mass shootings are distracting from the real danger of guns in America. *Washington Post,* available at: https://www.washingtonpost.com/news/wonk/wp/2015/12/07/mass-shootings-are-distracting-from-the-real-danger-of-guns-in-america/.

Gurney, K.; Teproff, C.; Rabin, C. (2017). A year after young king's death, toll from gun violence continues. *Miami Herald,* available at: https://www.miamiherald.com/news/local/education/article133231504.html.

Harding, D.J.; Fox, C.; Mehta, J.D. (2002). Studying rare events through qualitative case studies: lessons from a study of rampage school shootings. *Social Methods Res, 31*: 174-217.

Hart, T.C.; Lersch, K.M. (2015). *Space, time, and crime*. Carolina Academic Press: Durham, North Carolina.

Hemenway, D.; Miller, M. (2000). Firearm availability and homicide rates across 26 high-income countries. *The Journal of Trauma, Injury, Infection, and Critical Care, 49(6)*: 985–988.

Hempel, A. G.; Richards, T. C. (1999). Offender and offense characteristics of a nonrandom sample of mass murderers. *Journal of the American Academy of Psychiatry and the Law Online, 27(2):* 213–225.

Hepburn, L.M.; Hemenway, D. (2004). Firearm availability and homicide: A review of the literature. *Aggression and Violent Behavior, 9(4)*: 417-440.

Hepburn, L.; Miller, M.; Azrael, D.; Hemenway, D. (2007). The US gun stock: Results from the 2004 national firearms survey. Injury Prevention, 13(1): 15-19, doi: 10.1136/ip.2006.013607.

Hilal, S.M.; Densley, J.; Li, S.D.; Ma, Y. (2014). The routine of mass murder in China. *Homicide Studies, 18*: 83–104.

Ho Shon, P.; Roberts, M. (2010). An archival exploration of homicide-suicide and mass murder in the context of 19th-century American parricides. *International Journal of Offender Therapy and Comparative Criminology, 54:* 43–60. doi: 10.1177/0306624X08324472.

Hobbs, T.D. (2018). Most guns used in school shootings come from home. *Wall Street Journal*. Available at https://www.wsj.com/articles/in-school-shootings-most-guns-come-from-home-1522920600.

Holmes, R.M.; Holmes, S.T. (1992). Understanding mass murder: A starting point. *Federal Probation, 56:* 53-60.

Huff-Corzine, L.; McCutcheon, J.; Corzine, J. (2014). Shooting for accuracy: comparing data sources on mass murder. *Homicide Studies, 18*: 105–24. doi:10.1177/1088767913512205.

Humphreys, D.K.; Gasparrini, A.; Wiebe, D.J. (2016). Evaluating the impact of Florida's "Stand Your Ground" self-defense law on homicide and suicide by firearm: An interrupted time series study. *Journal of the American Medical Association*, doi: 10.1001/jamainternmed.2016.6811.

Infoplease. (2019). Timeline of worldwide school and mass shootings, available at: https://www.infoplease.com/history/world/timeline-of-worldwide-school-and-mass-shootings.

Isenstein, L. (2015). The states with the most gun laws see the fewest gun-related deaths. *The Atlantic*. Available at https://www.theatlantic.com/politics/archive/2015/08/the-states-with-the-most-gun-laws-see-the-fewest-gun-related-deaths/448044/.

Jager, E.; Goralnick, E.; McCarty, J.C.; Hashmi, Z.G.; Jarman, M.P.; Haider, A.H. (2018). Lethality of civilian active shooter incidents with and without semiautomatic rifles in the United States. JAMA, 320(10): 1034-1035, doi: doi:10.1001/jama.2018.11009.

Jeltsen, M. (2017). We're missing the big picture on mass shootings. *Huffington Post*. Available at http://www.huffingtonpost.com/entry/mass-shootings-domestic-violence-women_us_55d3806ce4b07addcb44542a.

Jetter, M.; Walker, J.K. (2018). The effect of media coverage on mass shootings. *IZA Institute of Labor Economics*, available at http://ftp.iza.org/dp11900.pdf.

Johnson, S.D.; Bernasco, W.; Bowers, K.J.; Elffers, H.; Ratcliffe, J.; Rengert, G. (2007). Space-time patterns of risk: A cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology, 23(3)*: 201–219, doi: 10.1007/s10940-007-9025-3.

Jung, R.S.; Jason, L.A. (1988). Firearm violence and the effects of gun control legislation. *American Journal of Community Psychology, 16*: 515-524.

Kahneman, D.; Tversky, A. (1973). On the psychology of prediction. *Psychol. Rev. 80*: 237–251.

Kalesan, B.; Mobily, M.E.; Keiser, O.; Fagan, J.A.; Galea, S. (2016). Firearm legislation and firearm mortality in the USA: a cross-sectional, state-level study. Lancet, 387:1847-1855, doi: 10.1016/S0140-6736(15)01026-0.

Kalesan, B.; Villarreal, M.D.; Keyes, K.M.; Galea, S. (2015). Gun ownership and social gun culture. *Injury Prevention*, 0: 1-5, doi: 10.1136/injuryprev-2015-041586.

Kalish, R.; Kimmel, M. (2010). Suicide by mass murder: Masculinity, aggrieved entitlement, and rampage school shootings. *Health Sociology Review, 19*: 451–464, doi: 10.5172/hesr.2010.19.4.451.

Kaplan, Jacob. Uniform Crime Reporting (UCR) Program Data: County-Level Detailed Arrest and Offense Data: county_ucr_offenses_known_1960_2017_rda.zip. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2019-02-10. https://doi.org/10.3886/E108164V3-10581.

Kaufman, E.J.; Morrison, C.N.; Branas, C.C.; Wiebe, D.J. (2018). State firearm laws and interstate firearm deaths from homicide and suicide in the United States: A cross-sectional analysis of data by county. *JAMA Intern Med, 178(5)*: 692-700, doi: 10.1001/jamainternmed.2018.0190.

Keneally, M. (2019). How new gun laws could stop some future would-be mass shooters. ABC News, available at: https://abcnews.go.com/US/current-gun-laws-stopped-deadliest-mass-shootings/story?id=63475147.

Kennedy-Kollar, D.; Charles, C. (2013). Hegemonic masculinity and mass murderers in the United States. *Southwest Journal of Criminal Justice, 8:* 62–74.

Kerjariwal, A. (2015). Introducing practical and robust anomaly detection in a time series. *Twitter*. Available at: https://blog.twitter.com/engineering/en_us/a/2015/introducing-practical-and-robust-anomaly-detection-in-a-time-series.html.

Killias, M. (1993). International correlations between gun ownership and rates of homicide and suicide. *Canadian Medical Association Journal, 148*: 1721-1725.

Killias, M.; Van Kesteren, J.; Rindlisbacher, M. (2001). Guns, violent crime, and suicide in 21 countries. *Canadian Journal of Criminology, 43*: 429-448.

King, G.; Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis, 9(2)*: 137-163.

Kissner, J. (2016). Are Active Shootings Temporally Contagious? An Empirical Assessment. Journal of Police and Criminal Psychology, 31(1), doi: https://doi.org/10.1007/s11896-015-9163-8.

Kivisto, A.J.; Ray, B.; Phalen, P.L. (2017). Firearm legislation and fatal police shootings in the United States. *Research and Practice,* doi: 10.2105/AJPH.2017.303770.

Kleck, G. (2016). Large-capacity magazines and the casualty counts in mass shootings: The plausibility of linkages. *Justice Research and Policy, 17(1)*: 28-47, doi: 10.1177/1525107116674926.

Kleck, G. (2004). Measures of gun ownership levels for macro-level crime and violence research. *Journal of Research in Crime and Delinquency 41(1):* 3-36.

Kleck, G. (2009). Mass shootings in schools. The worst possible case for gun control. *American Behavioral Scientist, 52(10):* 1447-64.

Knoll, J.L. (2010). The "pseudocommando" mass murderer: Part I, the psychology of revenge and obliteration. *Journal of the American Academy of Psychiatry and the Law Online, 38:* 87-94.

Knoll, J.L. (2012). Mass murder: Causes, classification, and prevention. *Psychiatric Clinics of North America, 35:* 757-780.

Koper, C.S. (2020). Assessing the potential to reduce deaths and injuries from mass shootings through restrictions on assault weapons and other high-capacity semiautomatic firearms. *Criminology and Public Policy, 19(1),* doi: https://doi.org/10.1111/1745-9133.12485.

Koper, C.S.; Johnson, W.D.; Nichols, J.L.; Ayers, A.; Mullins, N. (2017). Criminal use of assault weapons and high-capacity semiautomatic firearms: An updated examination of local and national sources. *Journal of Urban Health*, doi: 10.1007/s11524-017-0205-7.

Krouse, W.J. (2012). Gun control legislation. Congressional Research Service, Washington, DC. Available at https://fas.org/sgp/crs/misc/RL32842.pdf.

Krouse, W.J.; Richardson, D.J. (2015). Mass murder with firearms: Incidents and victims, 1999-2013. *Congressional Research Service.* Available at https://fas.org/sgp/crs/misc/R44126.pdf.

Kwon, I.W.; Bradle, S.; Safranski, S.R.; Bae, M. (1997). The effectiveness of gun control laws: Multivariate statistical analysis. *American Journal of Economics and Sociology, 56*: 41-50.

Kwon, R.; Cabrera, J.F. (2017). Social integration and mass shootings in U.S. counties. Journal of Crime and Justice, doi: 10.1080/0735648X.2018.1510336.

Kwon, R.; Cabrera, J.F. (2019). Income inequality and mass shootings in the United States. *BMC Public Health, 19(1147),* doi: Public Health (2019) 19:1147, https://doi.org/10.1186/s12889-019-7490-x.

LaFraniere, S.; Cohen, S.; Oppel Jr., R.A. (2015). How often do mass shootings occur? On average, every day, records show. *New York Times*. Available at http://www.nytimes.com/2015/12/03/us/how-often-do-mass-shootings-occur-on-average-every-day-records-show.html.

Langman, P.F. (2009). Rampage school shooters: A typology. *Aggression and Violent Behavior, 14*: 79–86, doi: 10.1016/j.avb.2008.10.003.

Langman, P.F. (2018). Role Models, Contagions, and Copycats: An Exploration of the Influence of Prior Killers on Subsequent Attacks. Available at: https://schoolshooters.info/sites/default/files/role_models_3.2.pdf.

Langman, P.F. (2016). School shooters: The warning signs, version 1.2, available at: https://schoolshooters.info/.

Lankford, A. (2012). A comparative analysis of suicide terrorists and rampage, workplace, and school shooters in the United States from 1990-2010. *Homicide Studies 17(3)*: 255-274.

Lankford, A. (2014). The myth of martyrdom: What really drives suicide bombers, rampage shooters, and other self-destructive killers. *Behavioral and Brain Sciences, 37(4)*: 351-393, doi: 10.1017/S0140525X13001581.

Lankford, A. (2015). Race and mass murder in the United States: A social and behavioral analysis. *Current Sociology, 64(3)*: 470–490, doi: 10.1177/0011392115617227.

Lankford, A. (2015b). Mass shooters in the USA, 1966–2010: Differences between attackers who live and die. *Justice Quarterly, 32*(2): 360-379, doi: 10.1080/07418825.2013.806675.

Lankford, A. (2016). Are America's public mass shooters unique? A comparative analysis of offenders in the United States and other countries. *International Journal of Comparative and Applied Criminal Justice, 40(2):* 1-13, doi: 10.1080/01924036.2015.1105144.

Lankford, A. (2016b). Public mass shooters and firearms: A cross-national study of 171 countries. Violence and Victims, *31(2),* 187-199, doi: 10.1891/0886-6708.

Lankford, A. (2016c). Detecting mental health problems and suicidal motives among terrorists and mass shooters. *Criminal Behaviour and Mental Health, 26*: 315-321, doi: 10.1002/cbm.2020.

Lankford, A.; Adkins, K.G.; Madfis, E. (2019). Are the Deadliest Mass Shootings Preventable? An Assessment of Leakage, Information Reported to Law Enforcement, and Firearms Acquisition Prior to Attacks in the United States. Journal of Contemporary Criminal Justice, 35(3), doi: https://doi.org/10.1177/1043986219840231.

Lankford, A.; Hakim, N. (2011). From Columbine to Palestine: A comparative analysis of rampage shooters in the United States and volunteer suicide bombers in the Middle East. *Aggression and Violent Behavior, 16*: 98–107, doi: 10.1016/j.avb.2010.12.006.

Lankford, A.; Madfis, E. (2018). Media Coverage of Mass Killers: Content, Consequences, and Solutions. American Behavioral Scientist, 62(2), doi: https://doi.org/10.1177/0002764218763476.

Lankford, A.; Silver, J. (2019). Why have public mass shootings become more deadly? Criminology and Public Policy, doi: https://doi.org/10.1111/1745-9133.12472.

Laqueur, H.S.; Wintemute, G.J. (2019). Identifying high-risk firearm owners to prevent mass violence. *Criminology and Public Policy, 19(1)*, doi: https://doi.org/10.1111/1745-9133.12477.

Larkin, R.W. (2009). The Columbine legacy: Rampage shootings as political acts. *American Behavioral Scientist, 52(9):* 1309–1326, doi: 10.1177/0002764209332548.

Lee L.K.; Fleegler, E.W.; Farrell, C.; Avakame, E.; Srinivasan, S.; Hemenway, D.; Monuteaux, M.C. (2016). Firearm laws and firearm homicides: A systematic review. *JAMA Internal Medicine*, doi: 10.1001/jamainternmed.2016.7051.

Lemieux, F. (2014). Effect of gun culture and firearm laws on gun violence and mass shootings in the United States: A multi-level quantitative analysis. *International Journal of Criminal Justice Sciences, 9(1)*: 74–93.

Lester, D. (2002). Trends in mass murder. *Psychological Reports*, 90: 1122, doi: 10.2466/pr0.2002.90.3c.1122.

Lester, D. (1993). Firearm deaths in the United States and gun availability. *American Journal of Public Health, 83(11)*: 1642.

Lester, D. (2010). Suicide in mass murderers and serial killers. *Suicidology Online, 1*: 19–27.

Lester, D.; Murrell, M.E. (1982). The preventive effect of strict gun control laws on suicide and homicide. *Suicide Life-Threat. Behavior, 12*: 131–140.

Lester, D.; Stack, S.; Schmidtke, A.; Schaller, S.; Muller, I. (2004). The deadliness of mass murderers. *Psychological Reports, 94*: 1404.

Lester, D.; Stack, S.; Schmidtke, A.; Scaller, S.; Muller, I. (2005). Mass homicide and suicide deadliness and outcome. Crisis: The Journal of Crisis Intervention and Suicide Prevention, 26(4): 184-187, doi: 10.1027/0227-5910.26.4.184.

Levin, J.; Madfis, E. (2009). Mass murder at school and cumulative strain a sequential model. *American Behavioral Scientist, 52(9):* 1227–1245, doi:10.1177/0002764209332543.

Lewiecki, E.M.; Miller, S.A. (2013). Suicide, guns, and public policy. *American Journal of Public Health, 103(1)*: 27-31, doi: 10.2105/AJPH.2012.300964.

Lewis, E.; Mohler, G.; Brantingham, P.J.; Bertozzi, A.L. (2011). Self-exciting point process models of civilian deaths in Iraq. *Security Journal, 25*: 244–264, doi: 10.1057/sj.2011.21.

Library of Congress (2015). Firearms-control legislation and policy: Australia. *The Law Library of Congress*, available at: http://www.loc.gov/law/help/firearms-control/australia.php.

Liwerant, O. S., 2007. Mass murder: Discussing criminological perspectives. *Journal of International Criminal Justice, 5:* 917-939.

Los Angeles Time. (2017). Deadliest U.S. mass shootings, 1984-2017, available at: https://timelines.latimes.com/deadliest-shooting-rampages/.

Lowe, S.R.; Galea, S. (2015). The mental health consequences of mass shootings. *Trauma, Violence, and Abuse, 18(1)*: 62-82, doi: 10.1177/1524838015591572.

Lu, Y.; Temple, J.R. (2019). Dangerous weapons or dangerous people? The temporal associations between gun violence and mental health. Preventive Medicine, 121, doi: https://doi.org/10.1016/j.ypmed.2019.01.008.

Luca, M.; Malhotraa, D.; Poliquina, C. (2017). Handgun waiting periods reduce gun deaths. *Proceedings of the National Academy of Sciences of the United States of America*, doi: 10.1073/pnas.1619896114.

Ludwig, J. (1998). Concealed gun carrying laws and violent crime: evidence from state panel data. *International Review of Law and Economics, 18*: 239-154.

Ludwig, J.; Cook, P.J. (2000). Homicide and suicide rates associated with implementation of the Brady Handgun Violence Prevention Act. *Journal of the American Medical Association, 284(5)*: 585–591.

Lynch, J., Jarvis, J. (2008). Missing data and imputation in the Uniform Crime Reports and the effects on national estimates. *Journal of Contemporary Criminal Justice, 24*: 69-85.

Maltz, M.D.; Targonski, J. (2002). A note on the use of county-level UCR data. *Journal of Quantitative Criminology, 18*: 297–318, doi: 10.1023/A:1016060020848.

Markowiak, S.F.; Heidt, D.G.; Welch, P.J. (2018). Social determinants of health identify communities at risk for mass shooting events. Journal of the American College of Surgeons, Scientific Forum Abstracts. Available at: https://www.facs.org/-/media/files/press-releases/abstracts2018/markowiak.ashx and https://www.facs.org/media/press-releases/2018/markowiak102318.

Mascia, J. (2019). In the Years Since the Isla Vista Shooting, the Incel Subculture Continues to Inspire Gunmen. *The Trace*. Available at: https://www.thetrace.org/2019/05/incel-anti-women-gun-violence-isla-vista-shooting/.

Mass Shooting Tracker (2019). Retrieved from https://www.massshootingtracker.org/data.

Masters, J. (2017). U.S. gun policy: Global comparisons. *Council on Foreign Relations*, available at: https://www.cfr.org/backgrounder/us-gun-policy-global-comparisons.

Mayors Against Illegal Guns. (2013). Analysis of recent mass shootings, available at: http://libcloud.s3.amazonaws.com/9/56/4/1242/1/analysis-of-recent-mass-shootings.pdf.

McDowall, D.; Loftin, C.; Wiersema, B. (1995). Easing concealed firearms laws: Effects on homicide in three states. *Journal of Criminal Law and Criminology, 86*: 193-206.

McPhedran, S.; Baker, J. (2011). Mass shootings in Australia and New Zealand: A descriptive study of incidence. *Justice Policy Journal, 8(1)*. Available at http://www.cjcj.org/uploads/cjcj/documents/Mass_shootings.pdf.

Meloy, J.R.; Hempel, A.G.; Mohandie, K.; Shiva, A.; Gray, B. (2001). Offender and offense characteristics of a nonrandom sample of adolescent mass murderers.

Journal of the American Academy of Child and Adolescent Psychiatry, 40(6): 719–728, doi: 10.1097/00004583-200106000-00018.

Meloy, J.R.; Hempel, A.G.; Gray, T.; Mohandie, K.; Shiva, A.; Richards, T. C. (2004). A comparative analysis of North American adolescent and adult mass murder. *Behavioral Sciences and the Law, 22*: 291-309, doi: 10.1002/bsl.586.

Messner, S.F.; Anselin, L.; Baller, R.D.; Hawkins, D.F.; Deane, G.; Tolnay, S.E. (1999). The spatial patterning of county homicide rates: an application of exploratory spatial data analysis. *Journal of Quantitative Criminology, 15(4)*: 423-450.

Miller, M.; Azrael, D.; Hemenway, D. (2002). Rates of household firearm ownership and homicide across U.S. regions and states, 1988-1997. *American Journal of Public Health, 92(12)*: 1988-1993.

Miller, M.; Azrael, D.; Hemenway, D. (2002b). Household firearm ownership and suicide rates in the United States. *Epidemiology, 13*: 517–524, doi: 10.1097/00001648-200209000-00006.pmid:12192220.

Miller, M.; Azrael, D.; Hepburn, L.; Hemenway, D.; Lippmann, S.J. (2006). The association between changes in household firearm ownership and rates of suicide in the united states, 1981-2002. *Injury Prevention, 12*: 178–182, doi: 10.1136/ip.2005.010850. pmid:16751449.

Miller, M.; Lippmann, S.J.; Azrael, D.; Hemenway, D. (2007). Household firearm ownership and rates of suicide across the 50 United States. *The Journal of Trauma: Injury, Infection, and Critical Care, 62(4)*: 1029-1035, doi: 10.1097/01.ta.0000198214.24056.40.

Mohler G.O.; Short, M.B.; Brantingham, P.J.; Schoenberg, F.P.; Tita, G.E. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association, 106(493)*, 100-108, doi: 10.1198/jasa.2011.ap09546.

Moorhouse, J.C.; Wanner, B. (2006). Does gun control reduce crime or does crime increase gun control? CATO J. 26: 103-124.

Mother Jones (2019). US mass shootings, 1982-2019: Data from Mother Jones' investigation. Available at http://www.motherjones.com/politics/2012/12/mass-shootings-mother-jones-full-data.

Mukherjee, R. (2020). How many mass shootings might have been prevented by stronger gun laws? *Los Angeles Times*, available at https://www.latimes.com/projects/if-gun-laws-were-enacted/.

Mulvey, E.P.; Cauffman, E. (2001). The inherent limits of predicting school violence. *American Psychologist, 56:* 797-802.

Nagin, D.S.; Koper, C.S.; Lum, C. (2020). Policy recommendations for countering mass shootings in the United States. *Criminology and Public Policy, 19*(1), doi: https://doi.org/10.1111/1745-9133.12484.

National Center for Juvenile Justice. (2018). EZASHR: Easy access to the FBI's supplemental homicide reports: 1980-2016. Available at: https://www.ojjdp.gov/ojstatbb/ezashr/asp/off_selection.asp.

Naval Postgraduate School. (2019). K-12 School Shooting Database. Available at https://www.chds.us/ssdb/.

Nekvasil, E.K.; Cornell, D.G. (2015). Prevalence and offense characteristics of multiple casualty homicides: Are schools at higher risk than other locations? *Psychology of Violence, 5(3)*: 236-245, doi: 10.1037/a0038967.

Newman, K.S. (2007). School rampage shootings. *Contexts, 6(2)*: 28-34, doi: 10.1525/ctx.2007.6.2.28.

Newman, K. S.; Fox, C. (2009). Repeat tragedy: Rampage shootings in American high school and college settings, 2002–2008. *American Behavioral Scientist, 52(9)*: 1286–1308, doi: 10.1177/0002764209332546.

Norton, K. (2019). Leahy report disputes gun violence, video game link. VT Digger, available at: https://vtdigger.org/2019/09/04/leahy-report-disputes-gun-violence-video-game-link/.

O'Toole, M.E. (2000). The school shooter: A threat assessment perspective. *Federal Bureau of Investigation, Critical Incident Response Group, National Center for the Analysis of Violent Crime*.

Overberg, P.; Upton, J.; Hoyer, M. (2013). USA Today research reveals flaws in mass-killing data. *USA Today,* available at: http://www.usatoday.com/story/news/nation/2013/12/03/fbi-mass-killing-data-inaccurate/3666953/.

Palermo, G.B. (1997). The Berserk syndrome: A review of mass murder. *Aggression and Violent Behavior, 2(1)*: 1-8.

Palermo, G.; Ross, L. (1999). Mass murder, suicide, and moral development: Can we separate the adults from the juveniles? *International Journal of Offender Therapy and Comparative Criminology, 43(1)*: 8-20, doi: 10.1177/0306624X99431002.

Papachristos, A.; Wildeman, C.; & Roberto, E. (2015). Tragic, but not random: The social contagion of nonfatal gunshot injuries. Social Science & Medicine, 125: 139-150, doi: 10.1016/j.socscimed.2014.01.056.

Parsons, C.; Weigend, E. (2016). America under fire: An analysis of gun violence in the United States and the link to weak gun laws. *Center for American Progress*. Available at https://cdn.americanprogress.org/wp-content/uploads/2016/10/11100940/AmericaUnderFire-report.pdf.

Paulsen, D.J.; Bair, S.; Helms, D. (2010). *Tactical crime analysis: Research and investigation*. CRC Press.

Paulsen, D.J.; Robinson, M.B. (2009). *Crime mapping and spatial aspects of crime*. Pearson.

Petee, T.A.; Padgett, K.G.; York, T.S. (1997). Debunking the stereotype: An examination of mass murder in public places. *Homicide Studies, 1(4):* 317-337.

Peters, R. (2013). Rational firearm regulation: Evidence-based gun laws in Australia. In: Webster DW, Vernick JS, eds. Reducing Gun Violence in America: Informing Policy With Evidence and Analysis, 195-204.

Peterson, J.; Densley, J. (2019). Op-Ed: We have studied every mass shooting since 1966. Here's what we've learned about the shooters. Los Angeles Times, available at: https://www.latimes.com/opinion/story/2019-08-04/el-paso-dayton-gilroy-mass-shooters-data.

Peterson, R.; Krivo, L. (2005). Macrostructural analyses of race, ethnicity, and violent crime: Recent lessons and new directions for research. *Annual Review of Sociology, 31*: 331-356, doi: 10.1146/annurev.soc.31.041304.122308.

Pew Research Center (2015). Views on gun laws unchanged after Aurora shooting. *Pew Research Center*. Available at http://www.people-press.org/2012/07/30/views-on-gun-laws-unchanged-after-aurora-shooting/.

Phillips, B.J. (2015). Deadlier in the U.S.? On Lone Wolves, Terrorist Groups, and Attack Lethality. Terrorism and Political Violence, 29(3), doi: https://doi.org/10.1080/09546553.2015.1054927.

Porter, M.D.; White,G. (2012). Self-exciting hurdle models for terrorist activity. *The Annals of Applied Statistics, 6(1):* 106–124, doi: 10.1214/11-AOAS513.

Presser, L. (2012). Getting on top through mass murder: Narrative, metaphor, and violence. *Crime Media Culture 8(1): 3-21*, doi:10.1177/1741659011430443.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rand Corporation. (2018). The Effects of Stand-Your-Ground Laws. Available at: https://www.rand.org/research/gun-policy/analysis/stand-your-ground.html.

Reeping, P.M.; Cerda, M.; Kalesan, B.; Wiebe, D.J.; Galea, S.; Branas, C.C. (2019). State gun laws, gun ownership, and mass shootings in the US: cross sectional time series. BMJ, 364, doi: https://doi.org/10.1136/bmj.l542.

Reuter, P.; Mouzos, J. (2003). Australia: A massive buyback of low-risk guns. Available at: http://faculty.publicpolicy.umd.edu/sites/default/files/reuter/files/gun%20chapter.pdf.

Roberts, P.; Fitzpatrick, K.M.; Smith, B.L.; Damphouse, K.R. (2013). Identifying the characteristics of communities where perpetrators live and pre-incident activity occurs. START National Consortium for the Study of Terrorism and Responses to Terrorism. Available at: https://www.start.umd.edu/sites/default/files/files/publications/START_IUSSD_FromExtremistToTerrorist_April2013.pdf.

Robinson, P.L.; Boscardin, W.J.; George, S.M.; Teklehaimanot, S.; Heslin, K.C.; Bluthenthal, R.N. (2009). The effect of urban street gang densities on small area homicide incidence in a large metropolitan county, 1994-2002. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 86(4), doi:10.1007/s11524-009-9343-x.

Rocque, M. (2012). Exploring school rampage shootings: Research, theory, and policy. *The Social Science Journal*, 49: 304-313, doi:10.1016/j.soscij.2011.11.001.

Rodríguez, A.A.; Hempstead, K. (2011). Gun control and suicide: The impact of state firearm regulations in the United States, 1995–2004. *Health Policy,* 101: 95–103, doi: 10.1016/j.healthpol.2010.10.005.

Rosengart, M.; Cummings, P.; Nathens, A.; Heagerty, P.; Maier, R.; Rivara, F. (2005). An evaluation of state firearm regulations and homicide and suicide death rates. *Injury Prevention, 11*: 77–83, doi: 10.1136/ip.2004.007062 pmid:15805435.

Rosner, B. (1983). Percentage points for a generalized ESD many-outlier procedure. *Technometrics, 25(2)*: 165-172.

Ruddell, R.; May, G.L. (2005). State background checks and firearms homicides. *Journal of Criminal Justice, 33(2)*: 127-136.

Rudolph, K.E.; Stuart, E.A.; Vernick, J.S.; Webster, D.W. (2015). Association between Connecticut's permit-to-purchase handgun law and homicides. *American Journal of Public Health, 105(8):* 49-54, doi: 10.2105/AJPH.2015.302703.

Salam, M.; & Stack, L. (2018). Do video games lead to mass shootings? Researchers say no. *New York Times*. Available at https://www.nytimes.com/2018/02/23/us/politics/trump-video-games-shootings.html.

Sampson,R.J.; Wilson, W.J. (1995). Toward a theory of race, crime, and urban inequality. In: Crime and Inequality, edited by John Hagan and Ruth D Peterson, 37-56. Stanford, CA: Stanford University Press.

Santaella-Tenorio, J.; Cerdá, M.; Villaveces, A.; Galea S. (2016). What do we know about the association between firearm legislation and firearm-related injuries? *Epidemiologic Reviews, 38(1)*: 140-157, doi: 10.1093/epirev/mxv012.

Schildkraut, J.H.; Elsass, H.J.; Meredith, K. (2017). Mass shootings and the media: why all events are not created equal. Journal of Crime and Justice, 41(3), doi: https://doi.org/10.1080/0735648X.2017.1284689.

Seitz, S.T. (1972). Firearms, homicide, and gun control effectiveness. *Law Soc Rev., 6*: 595–614.

Shankar, A. (2008). Spatial and temporal modeling of IED emplacements against dismounted patrols. Thesis, George Mason University. Available at http://ebot.gmu.edu/bitstream/handle/1920/9191/Shankar_gmu_0883E_10717.pdf.

Shaw, C.; McKay, H. (1942). *Juvenile delinquency and urban areas*. Chicago: University of Chicago Press.

Shenassa, E.D.; Daskalakis, C.; Buka, S.L. (2006). Utility of indices of gun availability in the community. Journal of Epidemiological Community Health, 60(1): 44-49, doi: 10.1136/jech.2005.039149.

Sherman, L.W. (2000). Gun carrying and homicide prevention. *Journal of the American Medical Association, 282:* 1193-1195.

Siegel M.; Ross, C.S.; King III, C. (2013). The relationship between gun ownership and firearm homicide rates in the united states, 1981–2010. *American Journal of Public Health, 103(11)*: 2098–2105, doi: 10.2105/AJPH.2013.301409.pmid:24028252

Siegel, M; Ross, C.S.; King, C. (2014). A new proxy measure for state-level gun ownership in studies of firearm injury prevention. *Injury Prevention, 20(3)*, 204–207.

Siegel, M.; Rothman, E.F. (2016). Firearm ownership and suicide rates among US men and women, 1981-2013. *American Journal of Public Health, 106(7)*, doi: 10.2105/AJPH.2016.303182.

Siegel, M.D.; Ziming, X.; Ross, C.S.; Galea, S.; Kalesan, B.; Fleegler, E.; Goss, K.A. (2017). Easiness of legal access to concealed firearm permits and homicide rates in the United States. *American Journal of Public Health, 107(12),* 1923-1929, doi: 10.2105/AJPH.2017.304057.

Silva, J.R.; Capellan, J.A. (2018). The media's coverage of mass public shootings in America: fifty years of newsworthiness. International Journal of Comparative and Applied Criminal Justice, 43(1), doi: https://doi.org/10.1080/01924036.2018.1437458.

Silva, J.R.; Greene-Colozzi, E.A. (2019). Fame-seeking mass shooters in America: Severity, characteristics, and media coverage. Aggression and Violent Behavior, 48, doi: https://doi.org/10.1016/j.avb.2019.07.005.

Silver, J. (2020). Space between concern and crime: Two recommendations for promoting the adoption of the threat assessment model and encouraging bystander reporting. *Criminology and Public Policy, 19*(1), doi: https://doi.org/10.1111/1745-9133.12474.

Silver, J.; Simons, A.; Craun, S. (2018). A study of pre-attack behaviors of active shooters in the United States between 2000 and 2013. *Federal Bureau of Investigation*. Available at https://www.fbi.gov/file-repository/pre-attack-behaviors-of-active-shooters-in-us-2000-2013.pdf/view.

Simonetti, J.A.; Rowhani-Rahbar, A.; Mills, B.; Young, B.; Rivara, F.P. (2015). State firearm legislation and nonfatal firearm injuries. American Journal of Public Health, 105: 1703-1709, doi: 10.2105/AJPH.2015.302617.

Skeem, J.L.; Mulvey, E. (2019). What role does serious mental illness play in mass shootings and how should we address it? *Criminology and Public Policy, 19(1),* doi: https://doi.org/10.1111/1745-9133.12473.

Smith, B. (2008). A Look at Terrorist Behavior: How They Prepare, Where They Strike. NIJ Journal, 260, available at: https://www.ncjrs.gov/pdffiles1/nij/222900.pdf.

Smith, B.; Roberts, P.; Damphouse, K. (2012). Update on geospatial patterns of precursor behavior among terrorists. Report to Human Factors/Behavioral Sciences Division, DHS Science and Technology Directorate. College Park, MD: Start, 2012. Available at: https://www.start.umd.edu/sites/default/files/files/publications/GeospatialPatternsofPrecursorBehaviorAmongTerrorists_Nov2012Update.pdf.

Smith, D. (2015). Twitter's new R package for anomaly detection. *Revolutions*. Available at https://blog.revolutionanalytics.com/2015/01/twitters-new-r-package-for-anomaly-detection.html.

Stanford Geospatial Center. (2019). The mass shootings in American database. Available at: https://github.com/StanfordGeospatialCenter/MSA.

Studdert, D.; Wintemute, G.; Zhang, Y.; Rodden, J. (2017). Handgun acquisitions in California after two mass shootings. *Annals of Internal Medicine*. Available at http://med.stanford.edu/news/all-news/2017/05/california-handgun-sales-spiked-after-two-mass-shootings.html.

Swanson, J.W.; Norko, M.A.; Lin, H.; Alanis-Hirsch, K.; Frisman, L.K.; Baranoski, M.V.; Easter, M.M.; Robertson, A.G.; Swartz, M.S.; Bonnie, R.J. (2016). Implementation and effectiveness of Connecticut's risk-based gun removal law: Does it prevent suicides? *Law and Contemporary Problems.*

Swedler, D.I.; Simmons, M.M.; Dominici, F.; Hemenway, D. (2015). Firearm prevalence and homicides of law enforcement officers in the United States. *American Journal of Public Health, 105(10)*: 2042-2048, doi: 10.2105/AJPH.2015.302749.

Taleb, N.N. (2010). *The black swan: The impact of the highly improbable*. Random House.

Team Trace. (2017). A guide to understanding mass shootings in America. *The Trace*. Available at https://www.thetrace.org/2017/01/understanding-mass-shootings-in-america/.

Telep, C.W.; Mitchell, R.J.; Weisburd, D. (2014). How much time should the police spend at crime hot spots? Answers from a police agency directed randomized field trial in Sacramento, California. *Justice Quarterly, 31(5)*: 905-933, doi: 10.1080/07418825.2012.710645.

Tessler, R.A.; Mooney, S.J.; Witt, C.E. (2017). Use of firearms in terrorist attacks: Differences between the United States, Canada, Europe, Australia, and New Zealand. *JAMA Internal Medicine, 177(12)*: 1865-1868, doi: 10.1001/jamainternmed.2017.5723.

Thompson, S.; Kyle, K. (2005). Understanding mass school shootings: Links between personhood and power in the competitive school environment. *Journal of Primary Prevention, 26(5)*: 419-438, doi: 10.1007/s10935-005-0006-8.

Thorndike, R.L. (1953). Who belongs in the family? *Psychometrika, 18*: 267-276, doi: https://doi-org.mutex.gmu.edu/10.1007/BF02289263.

Toffler Associates. (2011). Attack the network lexicon. Available at https://info.publicintelligence.net/JIEDDO-ATN-Lexicon.pdf.

Tonso, K.L. (2009). Violent masculinities as tropes for school shooters: The Montreal massacre, the Columbine attack, and rethinking schools. *American Behavioral Scientist, 52(9)*: 1266-1285, doi: 10.1177/0002764209332545.

Towers, S.; Gomez-Lievano, A.; Khan, M.; Mubayi, A.; Castillo-Chavez, C. (2015). Contagion in mass killings and school shootings. PLOS ONE, doi: 10.1371/journal.pone.0117259.

Tung, E.L.; Hawkley, L.C.; Cagney, K.A.; Peek, M.E. (2019). Social Isolation, Loneliness, And Violence Exposure In Urban Adults. Health Affairs, 38(10), doi: https://doi.org/10.1377/hlthaff.2019.00563.

Twemlow, S.W.; Fonagy, P.; Sacco, F.C.; O'Toole, M.E.; Vernberg, E. (2002). Premeditated mass shootings in schools: Threat assessment. Journal of the American Academy of Child and Adolescent Psychiatry, 41(4): 475-477.

United States Census Bureau [Census]. (2012). Geographic terms and concepts - census tract. Available at: https://www.census.gov/geo/reference/gtc/gtc_ct.html.

United States Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Multiple Cause of Death 1999-2018 on CDC WONDER Online Database, released 2020.

University of Maryland National Consortium for the Study of Terrorism and Responses to Terrorism. (2019). Global Terrorism Database. Available at: https://www.start.umd.edu/gtd/.

USA Today. (2019). Behind the bloodshed: The untold story of American's mass killings. Available at: http://www.gannett-cdn.com/GDContent/mass-killings/index.html#title.

Vasilogambros, M. (2018). After Parkland, States Pass 50 New Gun-Control Laws. Pew, available at: https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2018/08/02/after-parkland-states-pass-50-new-gun-control-laws.

Vecino-Ortiz, A.; Guzman-Tordecilla, D.N. (2019). Gun-carrying restrictions and gun-related mortality, Colombia: a difference-in-difference design with fixed effects. *Bulletin of the World Health Organization* 2020;98:170-176. doi: http://dx.doi.org/10.2471/BLT.19.236646.

Vigdor, E.R.; Mercy, J.A. (2006). Do laws restricting access to firearms by domestic violence offenders prevent intimate partner homicide? *Evaluation Review, 30(3)*: 313–346.

Violence Policy Center. (2020). Large capacity ammunition magazines, available at: http://www.vpc.org/fact_sht/VPCshootinglist.pdf.

Violence Policy Center. (2015). Mass shootings involving concealed handgun permit holders, available at: http://concealedcarrykillers.org/wp-content/uploads/2015/04/FACTSHEET-CCW-Mass-Shooters.pdf.

Vittes, K.A.; Vernick, J.S.; Webster, D.W. (2012). Legal status and source of offenders' firearms in states with the least stringent criteria for gun ownership. *Injury Prevention*, doi: 10.1136 /injuryprev-2011-040290.

Wagstaff, K.; Cardie, C.; Rogers, S.; Schroedl, S. (2001). Constrained K-means clustering with background knowledge. *Proceedings of the Eighteenth International Conference on Machine Learning, 2001, p. 577–584.* Available at: https://web.cse.msu.edu/~cse802/notes/ConstrainedKmeans.pdf.

Webster D.; Crifasi C.K.; Vernick, J.S. (2014). Effects of the repeal of Missouri's handgun purchaser licensing law on homicides. *Journal of Urban Health, 91(2)*: 293-302.

Webster, D.W.; Donohue, J.J.; Klarevas, L.; Crifasi, C.K.; Vernick, J.S.; Jernigan, D. (2016). Firearms on college campuses: Research evidence and policy implications. Baltimore Johns Hopkins Bloomberg School of Public Health.

Webster, D.W.; McCourt, A.D.; Crifasi, C.K.; Booty, M.D.; Stuart, E.A. (2020). Evidence concerning the regulation of firearms designed, sale, and carrying on fatal mass shootings in the United States. *Criminology and Public Policy, 19*(1), doi: https://doi.org/10.1111/1745-9133.12487.

Webster, D.W.; Vernick, J.S.; Hepburn, L.M. (2002). Effects of Maryland's law banning "Saturday Night Special" handguns on homicides. *American Journal of Epidemiology, 155(5)*: 406-412.

Webster, D.W.; Wintemute, G.J. (2015). Effects of policies designed to keep firearms from high-risk individuals. *Annual Review of Public Health, 36:* 21-37, doi: 10.1146/annurev-publhealth-031914-122516.

Wellford, C.F.; Pepper, J.V.; Petrie, C.V. (Eds.) (2004). *Firearms and violence: A critical review.* National Academies Press.

Wintemute, G.J. (2015). The epidemiology of firearm violence in the twenty-first century United States. *Annual Review of Public Health*, 36: 5-19.

Wintemute, G.J.; Pear, V.A.; Schleimer, J.P.; Pallin, R.; Sohl, S.; Kravitz-Wirtz, N.; Tomsich, E.A. (2019). Extreme risk protection orders intended to prevent mass shootings: A case series. *Annals of Internal Medicine,* doi: 10.7326/M19-2162.

Zeoli, A.M.; Paruk, J.K. (2019). Potential to prevent mass shootings through domestic violence firearm restrictions. *Criminology and Public Policy, 19(1),* doi: https://doi.org/10.1111/1745-9133.12475.

Zeoli, A.M.; Pizarro, J.M.; Grady, S.C.; Melde, C. (2014). Homicide as infectious disease: using public health methods to investigate the diffusion of homicide. *Justice Quarterly, 31(3)*: 609-632.

Zimring, F.E. (1968). Is gun control likely to reduce violent killings? *The University of Chicago Law Review, 35*: 721-737.

BIOGRAPHY

Matthew D'Anna graduated from Sandy Hook Elementary School, Newtown, Connecticut, in 1993. He received his Bachelor of Arts degrees from Arizona State University in 2004, his Master of Arts from John Jay College in 2008, and his Master of Advanced Study from Arizona State University in 2008. He is currently employed as a Chief Scientist at Booz Allen Hamilton in Northern Virginia since 2014.