**Title**
Internet Search Data: A Better Proxy for Regional Suicidality



**Author**
Guerra, Christian

**Publication Date**
2019

UNIVERSITY OF CALIFORNIA
RIVERSIDE

# Internet Search Data: A Better Proxy for Regional Suicidality

Master of Arts

in

Sociology

by

Christian Guerra

June 2019

Thesis Committee:
Dr. Bruce G. Link, Chairperson
Dr. Augustine Kposowa
Dr. Richard Carpiano

The Thesis of Christian Guerra is approved:

Committee Chairperson

University of California, Riverside

ABSTRACT OF THE THESIS

Internet Search Data: A Better Proxy for Regional Suicidality

by

Christian Guerra

Master of Arts, Graduate Program in Sociology
University of California, Riverside, June 2019
Dr. Bruce G. Link, Chairperson

Analyses of Google internet searches, reveal deep and otherwise elusive sentiments that are all too frequently unobserved by traditional research methods. Internet users often search topics they would be unwilling to ask their closest and most intimate confidants about. Thoughts about suicide, depression and the means one might use to take one's own life are prominent examples – people may keep these thoughts very private but reveal them to google. Spring boarding from this central insight this paper seeks to determine whether suicide rates are higher in places where google searches about depression and suicide are common. Building on prior research, this project will assess the extent to which Google searches predict the 50-state suicide rate above and beyond traditional approaches. Using factor analysis three clusters of search terms were identified: suicide, depression, and gun-related. Suicidal and depressive search terms are

found to be significantly associated with the suicide rate over the 11-year period (2006-2016) even after controlling for divorce, unemployment, mental state, binge drinking, gun ownership, and yearly effects. These findings suggest that internet search data represent a, practical, and cost-effective method for studying suicide and other large-scale sociological phenomena. As fields such as public health, epidemiology, and psychiatry effectively incorporate internet and big data-driven methodologies, it is imperative that sociologists carefully consider its capabilities given its increasing utilization and potential.

TABLE OF CONTENTS


| | |
|---|---|
| INTRODUCTION | 1 |
| HYPOTHESES | 10 |
| METHODS | 11 |
| ANALYSIS | 21 |
| STATISTICAL EQUATION | 24 |
| RESULTS | 24 |
| ROBUSTNESS CHECKS | 32 |
| DISCUSSION | 33 |
| LIMITATIONS | 38 |
| CONCLUSION | 39 |
| REFERENCES | 40 |

LIST OF FIGURES

FIGURE 1 11
Simplified model of this analysis

FIGURE 2 12
The Google Trends score equation

FIGURE 3 13
Chart showing yearly frequency score of a search term

FIGURE 4 14
Map showing regional frequency score

FIGURE 5 14-15
Chart showing frequency score and interest ranking

LIST OF TABLES

TABLE 1 21
Descriptive Statistics of the Variables

TABLE 2 26
Principal Factor Analysis Results of Google Search Terms with Orthogonal Varimax Rotation

TABLE 3 27
Fixed Effects Regression Results

TABLE 4 28
Fixed Effects Regression Results with Yearly Fixed Effects

TABLE 5 28
Fixed Effects Regression Results with Control Variables and Yearly Fixed Effects

TABLE 6 30
Fixed Effects Regression Results with Control Variables and Yearly Fixed Effects

TABLE 7 31
Fixed Effects Regression Results with Control Variables and Yearly Fixed Effects

Suicidology has seen abundant research within the last few decades. The fields of public health, epidemiology, psychiatry, and sociology have produced remarkable insights on the underlying causal pathways of our species' most personal act. Prominent pathways explored thus far include divorce (Neumayer 2003; Wray et al. 2011), unemployment (Neumayer 2003; Norstrom 1995), depression (Phillips et al. 2007), stress (Lester and Gunn 2016), alcohol abuse (Neumayer 2003; Norstrom 1995), and firearm ownership (Knopov et al. 2018; Kposowa et al. 2016; Vitt et al. 2018). However, at least to some extent, progress has been stalled in some domains of inquiry because researchers have had to rely on the self-report nature of various suicide risk factors including alcohol consumption, firearm ownership, and mental health state. Internet search data allows us to reveal more profound data in comparison to previous methods.

It is no secret that people lie. People lie about how often they exercise, how much they drink, how much fast food they eat, how they feel towards a political candidate, and how they really feel towards minorities. Survey methodologists acknowledge that answers to sensitive questions are often distorted by social desirability bias (Krumpal 2013). One of the most recent and notable instances where desirability bias had a major effect was the underestimation of Donald Trump's success during the 2016 presidential election. Many poll respondents may have been reluctant to admit their support for such a contentious figure. How can researchers make accurate predictions without making such egregious mistakes? The answer lies not in what people say, but rather in what they do. And most importantly, what they do in private, free from social repercussions.

*Why Google?*

Given the increasing popularity of the internet, more individuals actively seek information online regarding everything from where to buy dentures to their physical and mental health symptoms. In fact, the internet is the world's most relied on source for health information (Rice 2006). Symptoms associated with suicide are oftentimes mental, and thus, because of the stigmatizing nature of mental illness and the obstacles to care, individuals are even more likely to investigate their problems online (Ayers et al. 2013).

In addition to internet health queries, individuals often seek information regarding suicide and how to effectively kill themselves (Mok et al. 2016; Thornton et al. 2016). An estimated 54 percent of English-language websites that result from suicide-related search terms provide step-by-step information on effective suicide methods (Biddle et al. 2016). Information regarding suicide methods, detailed instructions, and materials are readily available online (Recupero et al. 2008). Thus, individuals who query suicide-related and suicide method-related terms may be at an increased risk of successful suicide.

Under the same logic, individuals feeling mental unease and symptoms of depression may be inclined to research these topics online, free from stigma. Given the links between depressive states and suicide (Phillips et al. 2007), individuals who query depression-related search terms may be at an increased risk of suicide. Lastly, the impulsive nature of suicide is well documented (Deisenhammer et al. 2009). Successful suicides are largely dependent on the effectiveness of its means and the access to such effective means. The most effective means responsible for half of all suicides in the U.S.

are firearms. Thus, areas in which people frequently query firearm-related terms, such as “gun and ammunition for sale,” may create a context in which access to effective means of suicide provides increased opportunities for successful suicide.

The world’s most utilized search engine, Google, provides a strong platform for observing the information-seeking activities of its users. Google handles roughly 2 trillion searches per year, 167 billion searches per month, and 5.5 billion searches per day (Sullivan 2016), making it the world’s largest data source available (Stephens-Davidowitz 2017). Nearly 80 percent of internet search queries in the U. S. are conducted through Google (Statista 2018). This much data provides us insight into the needs, wants, demands, and interests of its users (Jun et al. 2018). Eighty-seven percent of Americans have access to the internet, with most (77%) having immediate access via smartphone (Pew 2018). This easy access has facilitated heavy usage with the average American spending nearly nine hours a day online (Fox and Rainie 2014). Google search term data provides extensive evidence about what they are curious about, what they are thinking about, and what they are interested in. When this information is aggregated it gives us a partial window on the collective consciousness of our society.

Unfortunately, search engines do not produce user-identifiable individual information on who is conducting google searches that put them at risk of suicide, however, they do produce search term frequency data by time and geographic region. This analysis focuses on aggregate state-level characteristics that can be gleaned from Google searches and analyzes their effects on the suicide rate in states over time.

Because data are at the aggregate and not the individual level results from this analysis cannot be generalized to the individual level. However, as many public health services and firearm laws are directed from the state-level, the results of this analysis may inform state-based initiatives and policies that could reduce rates of suicide in those states.

*How Google Reveals*

Google search data reveals fascinating peculiarities on the human condition at any place and time. The power in these data lies in the fact that individuals query about subjects they dare not ask their fellow peers, neighbors, and families about. An individual eager to learn about their sexless marriage, mental health issues, deep insecurities, or prejudices are free to do so online without facing any social repercussions given the anonymity the internet confers. Thus, the internet provides access to information on topics that might be too embarrassing or uncomfortable to ask others. For this reason, Google search data, in many senses, is a “digital truth serum” that provides us with a substantial amount of information that would otherwise never be admitted to anybody (Stephens-Davidowitz 2017).

Google search data provides us with a new medium of observing human behavior that reveals what traditional methods often miss. For example, Google searches for terms such as “my mom beat me” and “my dad hit me” increased during the Great Recession and were closely tracked with the unemployment rate (Stephens-Davidowitz 2017). Every percentage point increase in the unemployment rate was associated with a three percent increase in the search rate for child abuse terms. Given that an overwhelming amount of child abuse cases go unreported, and social service agencies were understaffed because of the recession, reports of child abuse failed to observe this disturbing trend. Google searches, on the other hand, were able to shed light on this difficult-to-observe phenomenon.

Additionally, searches looking for off-the-books ways to terminate a pregnancy are higher in states that have passed laws restricting abortions (Stephens-Davidowitz 2017). Thus, Google search data shows us, in real-time, the unfortunate and painful realities many individuals find themselves in. This data provides insight into social phenomena that goes largely unbeknownst if traditional methods are all that is available. In a similar fashion, may we be able to observe an area's inclination towards suicide by observing search terms related to one's agonizing struggle of wanting to end their own life?

Given the findings Google searches provide, this research argues that Google searches will add to what we can learn from more traditional approaches and the variables they bring to bear. Thus, as social researchers our focus should be not only be on what people *say* they are doing, but rather to an examination of what they *are* doing; not only on what they *say* they think, but rather on what it *is* they think. Thus, the use of Google search data allows us to circumvent the social desirability bias and provides us with an important window into the human psyche. For the first time in human history are we able to observe what millions of individuals are unwilling to tell their closest and most intimate confidants. For this very reason, this paper argues that a state's inclination towards suicide can be measured by using Google search term data.

*Google and Suicide*

Thus far, Google data has been successful in describing and predicting human behavior in numerous fields. Ginsberg et al.'s 2009 study successfully tracking influenza outbreaks more effectively than the Center for Disease Control (CDC) helped establish search term methodologies in research on disease surveillance. Later that year, Google search term data helped predict the unemployment rate, and was found to have significant correlations to car and home sales (Choi and Varian 2009; 2012). Chae et al. (2015) found that areas that frequently Googled racially-charged terms were associated with an 8.2 percent increase in the black mortality rate in that area. Areas that queried racially-charged search terms more frequently were also found to be some of the worst-performing districts for Barack Obama during the 2012 presidential elections (Stephens-Davidowitz 2014). Thus, Google search term data has the potential to reveal meaningful social patterns and phenomena across geographic areas and can be applied to a wide range of academic fields.

The overall influence of the internet on suicide is not well known. Neither is the association between searching suicide-related terms online and the suicide rate. This study focuses on the United States, of which some studies have centered on, albeit not to the extent conducted in Japan, South Korea, and Taiwan. Using online questionnaires, Sueki 2012) finds that individuals in Japan who search for deliberate self-harm (DSH) terms on Google such as “self injury”, “wrist cutting”, and “overdose” were found to have significant lifetime suicidal behaviors compared to individuals who did not report searching for DHS terms. A time-series analysis in Taiwan reveals that increased Google searches for charcoal-burning suicide are positively associated with a rise in suicide by this method (Chang et al. 2015). Google searches for “carbon monoxide poisoning” are significantly associated with the frequency of suicides by gassing in Germany (Paul et al. 2017). Lastly, Google search terms “hydrogen sulfide”, “hydrogen sulfide suicide”, and “suicide hydrogen sulfide” are significantly associated with the incidence of suicide in Japan (Hagihara et al. 2011). The extent to which similar search terms have been studied in the United States calls for further research into suicide-method terms and their association to the 50-state suicide rate.

Like suicide-method search terms, suicide-related terms such as “suicide,” “suicidal,” and “suicide hotline” are found to be positively associated with the suicide rate (Arora et al. 2016). Engaging with suicide-related internet use is associated with higher levels of suicide ideation (Mok et al. 2015). In efforts to establish causality, suicide-related terms increased before an increase in the suicide rate in Japan (Hagihara et al. 2012). In addition to directly related terms, depression and stress-related search terms are positively associated with the suicide rate in South Korea (Song et al. 2014), the United Kingdom (Arora et al. 2016), and Japan (Sueki 2012). These findings are consistent with empirical research that links depression and serious psychological distress to an increased risk of suicide (Miller et al. 2007; Phillips et al. 2007). A study that issued a depression questionnaire triggered by depressive and suicidal search terms determined that half of all respondents (N: 4,709) reported a major depressive episode (Liu et al. 2014), thus linking individuals searching for depressive terms to depression. Alarmingly, over 75 percent of respondents reported to have never sought help. In large part, the internet served to provide the respondents with information regarding their symptoms.

*Firearms and Suicide*

Lastly, studies consistently find a positive association between measures of firearm prevalence and firearm suicide in the U.S. (Knopov et al. 2018; Kposowa et al. 2016; Miller et al. 2012; Vitt et al. 2018). Firearm accessibility and having a firearm in the home is positively associated with higher firearm suicide rates (Mann and Michel 2016; Wiebe 2003). These findings are consistent with research demonstrating the impulsive nature of suicide-attempts as half of suicide-survivors reported acting on their decision within ten minutes (Deisenhammer et al. 2009). Vitt et al. (2018) finds that an increase in the number of firearm background checks within a state causes a significant increase in the rate of firearm suicides. Previous search term research on suicide fails to address this critical factor. In other words, does the frequency of querying firearm-related terms increase that area's propensity towards suicide? As half of all suicides in the U.S. are committed via firearm (CDC 2018), this study aims to fill the gaps by incorporating this crucial risk factor through a digital lens.

*This Analysis*

Additional suicide risk factors identified in the suicidality literature include divorce (Neumayer 2003; Wray et al. 2011), unemployment (Neumayer 2003; Norstrom 1995), mental health state (Phillips et al. 2007), and binge drinking (Neumayer 2003; Norstrom 1995). Of the studies in the literature, none were able to include all these control variables. Instead, most of these studies are bivariate, focusing strictly on the association between Google searches and suicide. Very few include control variables that are limited to mental health state, divorce, and unemployment. Rarely do such studies incorporate a comprehensive and robust number of control variables. To add to the existing literature and to ensure model robustness, this analysis will control for such variables. Furthermore, this is the first study on the association between Google search terms and suicidality to include binge drinking and firearm ownership control variables.

Previous research has only focused on single-category search terms and used a limited amount of terms. It therefore may be very likely that a factor consisting of various terms may better assess the rate of suicide than only using single-term analyses. Single-term analyses that only use terms such as "suicide", may be failing to capture more embedded social patterns that emerge only when multiple search terms are used. Capturing the frequency of a single-term fails to capture what additional terms are correlated with it. For example, a single-term analysis of the term "suicide" may capture searches that are related to a celebrity suicide or a popular media reference. On the other hand, a multiple-term factor allows the analysis to focus on areas that not only search for “suicide,” but additionally search for “suicide hotline” as well. Theoretically, it makes more sense to concern ourselves with areas that frequently search certain combinations of search terms than to be focused only one search term.

For this reason, this analysis involves a factor analysis of an exhaustive list of search terms to more accurately specify which search terms are associated with higher rates of suicide. This methodology recognizes the complexity in assessing suicide outcomes by arguing that assessing suicide outcomes is not simply observing the frequency of one search term but necessitates an assessment of multiple search terms. In the same way, we argue areas that frequently search for “suicide” are less inclined towards suicide than areas that frequently search for not only “suicide,” but also “suicide hotline” and “cyanide.” Certainly, using the latter method would provide a much better assessment of suicide outcomes by accounting for the compounding risk factors. This analysis fills this critical limitation of past research by incorporating multiple-search term factors and is the first analysis to the author’s knowledge to do so. By doing so, this analysis strives to further elucidate underlying mechanisms that increase the risk of suicide in the U.S.

HYPOTHESES

H1: There is a positive association between suicide-related search terms and the suicide rate after controlling for divorce, unemployment, mental health state, binge drinking, and firearm ownership.

H2: There is a positive association between depression-related search terms and the suicide rate after controlling for divorce, unemployment, mental health state, binge drinking, and firearm ownership.

H3: There is a positive association between firearm-related search terms and the suicide rate after controlling for divorce, unemployment, mental health state, binge drinking, and firearm ownership.

METHODS

This study creates a data structure that includes information on state suicide rates, previously employed predictors and multi-term google-derived measures of suicidality, depression, and firearm-interest. This will allow a focus on the association between suicide, depression, and firearm-related search terms and the state suicide rate for an eleven-year period (2006-2016) in the United States.

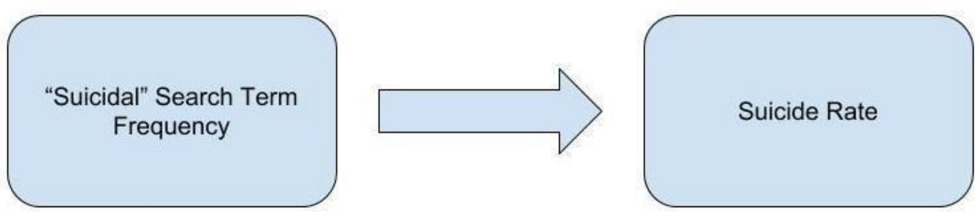


Figure 1. Independent variable is the frequency of suicidal Google searches and the dependent variable is the state suicide rate for that corresponding year.

*Independent Variable*

The independent variable is the frequency of suicide, depression, and firearm-related Google search terms per state during the years 2006 through 2016. Data on search terms are obtained from Google Trends. Google Trends is a marketing tool offered by Google that provides data on how often a search term is entered *relative to the total search volume* across that area. Google Trends provides a numerical variable ranging from 0 to 100 as described below. A score of 100 indicates the state that has the highest search frequency of a given term relative to the total of all other searches in that state. Yearly Google search term frequency data allows a user to observe as far back as 2006.

The numerical value provided by Google Trends indicates how frequently the term is searched for relative to *the total of all other terms searched for* in the state. This method ensures that areas with the highest populations will not be given the highest scores. The state that searches for “suicide” for example, in the greatest proportion relative to the total of all searches, will be assigned the score of 100. Thereon, the remaining states are assigned a value that represents that state’s search frequency relative to the highest scoring state. A simplified equation demonstrating the Google Trends score methodology can be observed below.

$$\frac{\textbf{(search term: “suicide”)}}{\textbf{(total number of all search terms in state)}} \quad \textbf{====} \quad \textbf{Google Trends score (0 - 100)}$$

Figure 2. The Google Trends score equation

To further elucidate the functionality of the Google Trends search tool, we type in the search term "mittens" into the search bar and choose the year "2017." From there, two illustrations and one dataset are produced. Pictured below (Figure 3) is the "Interest Over Time" chart which shows the relative search term volume of "mittens" over the year 2017. We can observe a negligible amount of interest in "mittens" during the Spring and Summer months. However, by October we observe a rising trend in searches for "mittens" followed by a precipitous rise in November through January, coinciding with the year's coldest weather.

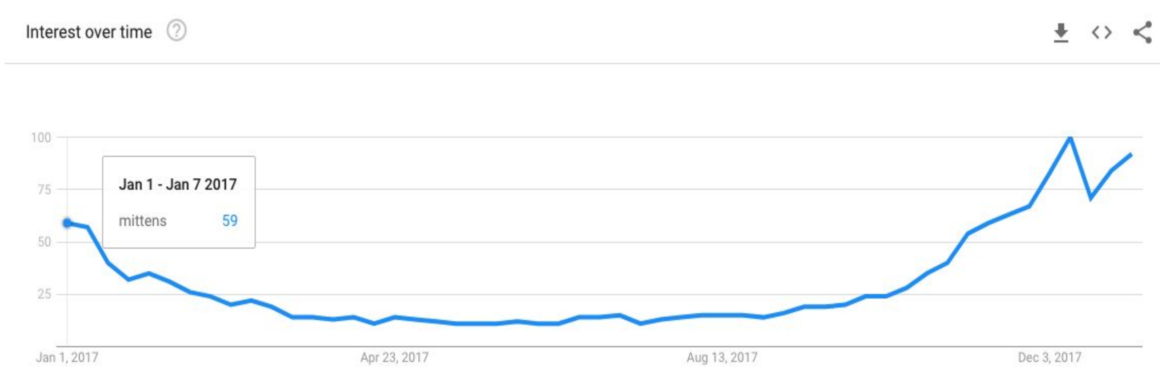


Figure 3. Interest of search term "mittens" over a one-year period. (January 2017 - December 2017)

Next, we observe the "Interest by Subregion" illustration provided by Google Trends (Figure 4). In this instance, the subregions shown are states. States in which individuals Google-search for "mittens" at the highest rates are shaded darker whereas states that neglect this winter accessory are shaded lighter. Consistent with the colder weather observed in the northern and northeastern states, people in these states Google for “mittens” at the highest rates as a proportion of all searches in the state. Additionally, we also can observe the southern states relative indifference for “mittens.”

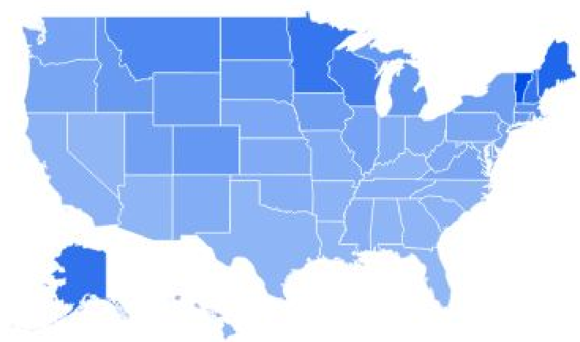

Figure 4. Interest of search term “mittens” by state. Dark blue indicates higher search term frequency. Light blue indicates lower search term frequency.

Lastly, Google Trends provides a quantitative dataset of all 50 states, and their search interest score. Consistent with the previous illustrations, we observe that the highest scoring states have severely cold weather during the winter months (Figure 5). Conversely, the lowest scoring states are ones where the weather is known to be warmer.


Figure 5. State interest score of search term “mittens.” State ranking on the left. Google Trends score on the right.

Search terms are guided by probabilistic, theoretical, and empirical postulations of key words an individual contemplating suicide is most likely to search for to gain familiarity with suicide methods. Additionally, the terms used in this study are found frequently enough on Google such that complete state data is provided. Often, when a search term is not frequently searched for there is not enough data for scores to be calculated across states and times.

Thus far, this study has gathered the following terms to coincide with symptoms of suicidal behavior and suicide ideation, which may be, in turn associated with the act itself:

Suicide, Suicidal, Commit Suicide, Kill Myself, Want to Die, Kill Yourself, Suicide Hotline, Suicide Prevention, Suicidal Thoughts, How to Overdose, How to Die, How to Suicide, Suicide By, Suicide Death, Assisted Suicide, Ways to Die, Self Harm, Cyanide, Carbon Monoxide, Carbon Monoxide Poisoning, Hydrogen Sulfide, Depression, Depressed, Depressing, Depression Test, Depression Help, Depression Quotes, Depressing Quotes, Signs of Depression, Symptoms of Depression, Gun, Guns for Sale, Buy Guns, Firearm, Ammo, Ammo for Sale, Shotgun, and Rifle.

The following terms were considered but failed to provide complete data:

Depression Suicide, Suicide Method, Suicide Forum, Suicide Reddit, How to Commit Suicide, Why Commit Suicide, Ways to Commit Suicide, I Want to Die, Suicide by Hanging, Suicide by Gun, Gun Suicide, Death Suicide, Death by Hanging, and Suicide Pills.

A factor analysis of all terms in every state for the year 2016 managed to separate the terms into three distinct factors: suicide, depression, and firearm-related. The factors are comprised of the terms with the highest loadings on the factor. With respect to suicide-related terms the searches with the highest loadings are comprised of the following: “suicide”, “suicidal”, “suicide hotline”, “suicide prevention”, and “cyanide.” Depression-related terms are comprised of the following searches: “depression”, “depressed”, “depression help”, and “kill yourself.” Firearm-related terms are comprised of the following searches: “gun”, “guns for sale”, “ammo”, “rifle”, and “shotgun.” Further details and results of the factor analysis are found under the results section.

The final variable comprises of the Google Trends scores for each term within each factor divided by the number of terms. For example, the suicide-terms variable is comprised of the following: (“suicide” score + “suicidal” score + “suicide hotline” score + “suicide prevention” score + “cyanide” score) / 5 = Suicide terms variable. The same methodology was applied to the depression and firearm factors. All factors continue to be on a scale from 0 to 100.

*Dependent Variable*

The dependent variable is the state suicide rate per 100,000 during the years 2006 through 2016. Suicide is defined as intentional self harm and defined by ICD-10 cause codes X60-X84 and Y87. Suicide rate data is obtained from the Center for Disease Control's WISQARS dataset. The data provides suicide rate figures per 100,000 inhabitants for each year for all 50 U.S. states. The data was moderately skewed, however not to the extent that it warranted a log transformation. Several studies argue for Poisson and negative binomial regression models when variables involve count data, such as suicide rates (Moksony and Hegedus 2014). However, these models work most effectively when count variables contain large amounts of zero entries, thus skewing the distribution. This was not the case with the suicide data which showed a range of 6.36 (New Jersey in 2009) to 29.65 (Wyoming in 2012). The variable's mean of 14.43 and standard deviation of 3.92 makes it appropriate for an OLS panel analysis.

*Control Variables*

Given the association between suicide and unemployment (Norstrom 1995), divorce (Wray et al. 2011), binge drinking (Norstrom 1995), mental health state (Phillips et al. 2007), and firearm ownership (Vitt et al. 2018), this analysis will control for all state-level variables to determine whether *Google search data measures suicide rates above and beyond these other known and well-established predictor variables*.

State-level divorce data from 2006 through 2016 was obtained from the CDC National Vital Statistics System. Several states (California, Georgia, Hawaii, Indiana, Louisiana, and Minnesota) do not provide divorce data. Missing data was replaced using mean imputation. The minimum value is 1.2 (Iowa in 2015) and the maximum value is 6.7 (Nevada in 2006). State-level unemployment data from 2006 through 2016 was obtained from the Bureau of Labor Statistics. No missing values are present in this data. Values range from 2.6 (Hawaii in 2006 and Utah in 2007) to 13.7 (Michigan in 2009).

Controlling for divorce and unemployment is consistent with the sociological literature that details disruptions and loss of familial and integrative ties, such as marriages and occupations, are associated with an increased risk of suicide. Those considered to be less integrated into society, such as the divorced, childless, single, unemployed, and occupationally redundant, are most at risk of suicide. Conversely, those considered more integrated, such as the married and employed, are protected from suicide by way of their social ties and integration.

State-level binge drinking data was obtained from the CDC Behavioral Risk Factor Surveillance System (BRFSS). The BRFSS is a nationally representative annual telephone survey conducted by the CDC that assesses health-related risk behaviors throughout the United States. At more than 400,000 adult interviews every year, it is the largest health survey system in the world (CDC 2019). The binge drinking variable is represented by the prevalence of respondents who consumed more than five drinks on an occasion in the last 30 days. This variable ranged from a minimum of 6.6 (Tennessee in 2010) to a maximum of 26.5 (Wisconsin in 2012).

State-level mental health state data was also obtained from the CDC BRFSS. Respondents were asked "Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?" The mental health state variable is represented by the mean number of days reported. This variable ranged from a minimum of 2.23 (North Dakota in 2008) to 5.23 (Alabama in 2012). Controlling for binge drinking and mental health state is consistent with the psychiatric literature that links such behaviors to suicide.

A methodological sampling discrepancy in the longitudinal BRFSS data that includes the binge drinking and mental health state variable occurred in 2011. Before 2011, the CDC's sampling methodology did not include mobile phone users. However, after 2011, the CDC began including mobile phone users into the survey. To control for this methodological change in the BRFSS's sampling methods, a dummy variable is introduced to the model. Thus, years 2006 through 2010 are represented by a value of 0, and years 2011 through 2016 are represented by a value of 1.

State firearm ownership is obtained from the Federal Bureau of Investigation's National Instant Criminal Background Check System (NICS). This background check system is used to ensure firearms are not sold to those with a history of criminal antecedents or those otherwise deemed ineligible. The FBI reports on the number of firearm background checks that are initiated through the NICS. States are ordered to conduct background checks for firearm sales through the NICS however, some states do not cooperate fully thus the data provides a somewhat incomplete estimate of firearm sales. The use of this proxy for firearm ownership was first used by Lang (2013) in a panel analysis demonstrating a significant positive association between firearm background checks and the firearm suicide rate.

The last year reliable firearm ownership data was obtained was in 2004 through the CDC's BRFSS. The BRFSS has since stopped asking respondents about firearm ownership. This has spurred researchers to attempt to obtain more reliable proxies for firearm ownership which include subscriptions to the magazine *Guns & Ammo* (Duggan 2003), using the firearm suicide rate (Azrael et al. 2004; Kleck 2004), and using the General Social Survey (GSS) (Miller et al. 2002). However, doubts have been cast on the magazine subscription proxy (Cook and Ludwig 2006), the firearm suicide rate measure has inherent simultaneity bias, and the GSS lacks statistical power (Lang 2013).

Given the difficulty of obtaining current and reliable firearm ownership data, the NICS provides a reliable proxy by reporting changes in the number of firearm owners, changes in the number of firearms in the state, and changes in the intent to purchase a firearm (Vitt et al. 2018). This study chose to focus on the handgun category of the dataset given that long guns such as rifles and shotguns are rarely used in suicides (Wintemute et al. 1988). The yearly total of handgun sales reported by the NICS in each state was gathered from years 2006 through 2016. The yearly totals were then divided by that state's yearly population then multiplied by 1,000 to produce rates. The states with some of the lowest rates of handgun sales reported by the NICS include New Jersey, Hawaii, New York, and Massachusetts. The states with some of the highest rates include Wyoming, Alabama, Alaska, West Virginia, and Montana.

**TABLE 1**

**Descriptive Statistics of the Variables**

| Variable (2006-2016) | Mean | Std. Dev. | Skew | Min. | Max. | N |
|---|---|---|---|---|---|---|
| Suicide Rate per 100,000 | 14.43 | 3.92 | 0.62 | 6.36 | 29.65 | 550 |
| Suicidal Terms Trends Score | 57.19 | 10.71 | 0.35 | 30.8 | 91.4 | 550 |
| Depression Terms Trends Score | 68.19 | 11.29 | -0.46 | 32.75 | 98.5 | 550 |
| Gun Terms Trends Score | 57.06 | 15.75 | 0.26 | 26.8 | 99.6 | 550 |
| Divorce Rate | 3.61 | 0.79 | 0.58 | 1.2 | 6.7 | 550 |
| Unemployment Rate | 6.27 | 2.17 | 0.66 | 2.6 | 13.7 | 550 |
| Binge Drinking Percentage | 16.66 | 3.5 | 0.03 | 6.6 | 26.5 | 550 |
| Mental Health State | 3.56 | 0.52 | 0.16 | 2.23 | 5.23 | 550 |
| Handgun Ownership Rate per 1,000 | 18.78 | 12.25 | 0.42 | 0 | 51.05 | 550 |
| BRFSS Dummy | 0.55 | 0.5 | -0.18 | 0 | 1 | 550 |

## ANALYSIS

This study will observe yearly (2006-2016) associations between suicidal Google search terms and the state suicide rate. A panel analysis of significant trends in search term rates and suicide rates over the 2006-2016 period is conducted via fixed effects modeling.

*On collinearity*

The presence of multicollinearity was assessed by calculating variance inflation factors for all independent variables under all three separate search term models. Variables producing values greater than 10 are considered to have excessive multicollinearity. The suicide-terms model produced a mean VIF of 1.71 and no variable exceeded 2.14 (BRFSS Dummy). The depression-terms model produced a mean VIF of 1.74 and no variable exceeded 2.46 (BRFSS Dummy). Lastly, the firearm-terms model produced a mean VIF of 1.78 and no variable exceeded 2.06 (BRFSS Dummy). Multicollinearity was not found to be a significant problem for the analyses presented below.

*Fixed Effects*

This analysis is estimated using a fixed effects model consisting of panel data for all 50 states for the period of 2006 through 2016. Fixed effects models allow us to control for potential omitted variable bias and unobserved time-invariant state-specific factors that may affect the rate of suicides. Variables such as cultural factors, climate, geography, elevation, religion, gender, and race are largely captured by the fixed effects as these variables rarely undergo drastic changes within states over time. Time-invariant differences between states, such as cultural norms and beliefs, may influence the suicide rate of such state. However, by focusing only on the within-state variation over time, fixed effects models can control for such exogenous variables.

In addition to controlling for time-invariant between-state differences, this model also controls for factors that may affect suicide trends at the national level by using yearly fixed effects. Numerous studies cite the effect that economic cycles have on suicide rates (Wasserman 1984). Time-specific social currents such as recessions or major events that can affect suicide rates on a national level can provide misleading estimates. In efforts to not confound the estimates with an exogenous variable that may affect national suicide trends, such as the Great Recession, the specified model will control for such trends using year fixed effects. Additionally, it may be the case that a popular movie or a celebrity suicide may spur national increases in suicidal searches and in the suicide rate, nevertheless such instances are controlled for by the yearly fixed effects.

Fixed effects models effectively control for all invariant between-state and yearly-trend differences. However, variables that vary from year to year could have an influence on the rate of suicide and therefore act as confounding variables on the relationship between Google searches and suicide rates. In accordance with the sociological literature that holds social and familial disintegration may lead suicide, this model controls for the state divorce rate and state unemployment rate. In accordance with the epidemiological literature that finds access to effective means of suicide increases the suicide rate, this model controls for the state handgun ownership rate. Lastly, as suicidology and psychiatry literatures have established the link between depressive moods and binge drinking to suicide, this model also controls for such. Model specification was conducted with a Hausman test. The results of the Hausman test indicate the appropriateness of the fixed effects over the random effects model.

STATISTICAL EQUATION

The multivariate fixed effects model to be estimated is as follows:

$$SR_{ij} = \beta ST_{it} + \beta X_{ij} + \gamma_t + \alpha_i + \varepsilon_{ij}$$

The term SR is the suicide rate per 100,000 in state *i* at time *t*. The variable of interest, search terms, is represented by vector ST. Vector X consists of the following time variant control variables: unemployment, divorce, mental health state, binge drinking, and handgun ownership. Parameter $\gamma$ represents time effects that account for unobserved national forces driving suicide rate. Parameter $\alpha$ represents unobserved state-specific fixed effects. The error term is represented by vector $\varepsilon$.

RESULTS

*Factor Analysis on Search Terms*

Given the large number of Google search terms a factor analysis was appropriate to determine which search terms reveal trend commonalities between terms. All forty terms from 2016 were correlated for observational purposes, examined through a factor analysis, rotated, separated by loadings, and then factor-derived scales were checked for internal consistency using Cronbach's alpha.

All forty terms were run through a factor test to ensure the terms were suitable for factor analysis. The factor test included a Bartlett test of sphericity and a Kaiser-Meyer-Olkin measure of sampling adequacy. The Bartlett test of sphericity produced a p-value of 0.000, thus rejecting the null hypothesis that the terms are not sufficiently intercorrelated enough to conduct a factor analysis. The Kaiser-Meyer-Olkin measure of sampling adequacy produced a value of 0.619. Values below .50 are considered not suitable for factor analysis. The value of .619 indicates that the variables are moderately suitable for factor analysis.

A principal factor analysis was then conducted on all forty terms. The factors were then rotated via orthogonal varimax rotation and sorted based on the magnitude of factor loadings. The factor analysis managed to separate the terms under three category factors: suicide-related, depression-related, and firearm-related. The factor analysis thus confirms the uniqueness of the separate terms which allowed for the construction of more robust variables. The top search terms with the heaviest and most unique loadings on each factor were then combined into new variables. These new variables with their respective search terms are as follows:

**Suicide Factor**: suicide, suicide prevention, cyanide, suicide hotline, suicidal

**Depression Factor**: depression, depression help, kill yourself, depressed

**Firearm Factor**: ammo, gun, shotgun, guns for sale, rifle

**TABLE 2**

**Principal Factor Analysis Results of Google Search Terms with Orthogonal Varimax Rotation**

| Search Term | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| Ammo | 0.93 | | |
| Gun | 0.89 | | |
| Shotgun | 0.87 | | |
| Guns for sale | 0.84 | | |
| Rifle | 0.80 | 0.50 | |
| Suicide | | 0.85 | |
| Suicide prevention | 0.34 | 0.85 | |
| Cyanide | | 0.77 | |
| Suicide hotline | | 0.73 | |
| Suicidal | | 0.57 | 0.39 |
| Depression | 0.33 | | 0.68 |
| Depression help | 0.36 | | 0.67 |
| Kill yourself | | | 0.64 |
| Depressed | | | 0.62 |
| | **α: 0.95** | **α: 0.89** | **α: 0.77** |

Loadings over 0.30 are shown. Alpha values were generated via Cronbach's Alpha.

The factor analysis results indicate strong internal consistency for the firearm and suicide factors (α: 0.95, α: 0.89 respectively) and moderate internal consistency for the depression factor (α: 0.77).

*Fixed Effects Results*

The analysis includes two separate models. The first model is a bivariate analysis of the search terms and the suicide rate. The second model adds the control variables that include divorce, unemployment, binge drinking, mental health state, firearm ownership, and the BRFSS dummy. The data was formatted into a panel dataset and run through a fixed effects model.

*Model 1*

The suicide, depression, and firearm factors were regressed individually on the suicide rate in a bivariate analysis.

**TABLE 3**
**Fixed Effects Regression Results**

| | β | *t* | *p* |
|---|---|---|---|
| **Google Suicide Terms** | .100*** | 10.69 | 0.000 |
| | (0.081 - 0.119) | | |
| **Google Depression Terms** | .088*** | 11.42 | 0.000 |
| | (0.072 - 0.103) | | |
| **Google Firearm Terms** | .107*** | 6.06 | 0.000 |
| | (0.071 - 0.143) | | |

*** indicates significant at the 1% level. ** indicates significant at the 5% level. * indicates significant at the 10% level. Robust confidence intervals in parenthesis.

The results of the fixed effects model demonstrate that all three factors are significantly associated with the suicide rate. The results indicate there is a significant association between the state suicide rate and suicide, depression, and firearm-related search terms in a bivariate analysis. When yearly fixed effects are implemented, more conservative results are produced. The yearly fixed effects captures much of the variance attributable to trends and thus, suicide-related searches remain significant however, depression and firearm-related search terms cease to be significant.

**TABLE 4**

**Fixed Effects Regression Results with Yearly Fixed Effects**

| | β | *t* | *p* |
|---|---|---|---|
| **Google Suicide Terms** | .025** | 2.20 | 0.032 |
| | (0.002 - 0.047) | | |
| **Google Depression Terms** | .016 | 1.55 | 0.128 |
| | (-0.005 - 0.038) | | |
| **Google Firearm Terms** | .006 | 0.42 | 0.675 |
| | (-0.022 - 0.033) | | |

*** indicates significant at the 1% level. ** indicates significant at the 5% level. * indicates significant at the 10% level. Robust confidence intervals in parenthesis.

*Model 2*

**TABLE 5**

**Fixed Effects Regression Results with Control Variables and Yearly Fixed Effects**

| | β | *t* | *p* |
|---|---|---|---|
| **Google Suicide Terms** | **.026**** | 2.34 | 0.023 |
| | (0.004 - 0.049) | | |
| **Divorce** | -.373 | -1.55 | 0.127 |
| | (-0.855 - 0.109) | | |
| **Unemployment** | .096 | 1.38 | 0.173 |
| | (-0.044 - 0.235) | | |
| **Binge** | .042 | 0.93 | 0.359 |
| | (-0.049 - 0.132) | | |
| **Mental State** | .109 | 0.40 | 0.690 |
| | (-0.437 - 0.655) | | |
| **Handgun Ownrshp** | **.027**** | 2.24 | 0.030 |
| | (0.003 - 0.051) | | |
| **BRFSS Dummy** | **1.80***** | 4.87 | 0.000 |
| | (1.059 - 2.547) | | |

Results are with state and year fixed effects. The dependent variable is the suicide rate per 100,000 population. Robust confidence intervals in parentheses. *** indicates significant at the 1% level. ** indicates significant at the 5% level. * indicates significant at the 10% level.

Model 2 includes the control variables of divorce, unemployment, mental health state, binge drinking, handgun ownership, the BRFSS dummy variable, and yearly time trends. The suicide terms factor is found to be significant at the five percent level. The handgun ownership is also found to be significant at the five percent level. The null hypothesis that suicide related terms has no association with the suicide rate can be rejected. The coefficient of the suicide factor on the suicide rate can be interpreted as every one unit increase of suicidal Google search term frequency in a state is associated with a 0.026 increase in the suicide rate of such state. A 0.026 increase in the suicide rate equates to approximately 83 suicides per year. Thus, every one unit increase in suicidal search term frequency is associated with 83 suicides in that year. This estimate is generated by dividing 0.026 by 100,000 then multiplying by 320,000,000 (average size of the US population in 2010).

On firearm ownership, every one unit increase in the rate of firearm ownership in a state is associated with 86 suicides in that state every year. That is, every 334,000 handgun purchase background checks in a state is associated with 86 suicides per year. In smaller terms, every 3,900 handgun background checks are associated with one suicide per year. This estimate is generated by dividing the mean of the firearm variable (18.8) by 1,000 then multiplying by 320,000,000 (average size of the US population in 2010) to estimate how many background checks take place on average every year. Then this figure is divided by 18.8 to produce the number of background checks needed to increase the rate of handgun ownership by one per 1,000.

**TABLE 6**

**Fixed Effects Regression Results with Control Variables and Yearly Fixed Effects**

| | β | *t* | *p* |
|---|---|---|---|
| **Google Depression Terms** | **.016*** | 1.73 | 0.090 |
| | (-0.003 - 0.034) | | |
| **Divorce** | -.315 | -1.24 | 0.222 |
| | (-0.828 - 0.197) | | |
| **Unemployment** | .054 | 0.95 | 0.345 |
| | (-0.059 - 0.167) | | |
| **Binge** | .039 | 0.86 | 0.393 |
| | (-0.052 - 0.130) | | |
| **Mental State** | .119 | 0.44 | 0.665 |
| | (-0.430 - 0.668) | | |
| **Handgun Ownrshp** | **.027**** | 2.25 | 0.029 |
| | (0.003 - 0.052) | | |
| **BRFSS Dummy** | **2.09***** | 5.05 | 0.000 |
| | (1.259 - 2.922) | | |

Results are with state and year fixed effects. The dependent variable is the suicide rate per 100,000 population. Robust confidence intervals in parentheses. *** indicates significant at the 1% level. ** indicates significant at the 5% level. * indicates significant at the 10% level.

The depression terms factor is significant at the ten percent level however, handgun ownership is significant at the five percent level. The null hypothesis that depression-related terms have no association with the suicide rate can be rejected, although at the ten percent level. Caution must be taken when interpreting the depression-term results as the probability of committing a type one error is high. The coefficient of the depression factor on the suicide rate can be interpreted as every one unit increase of depressive search term frequency in a state is associated with 51 suicides in that state per year. The firearm ownership proxy produces similar results to the previous model that tests suicide terms ( $\beta = 0.027$). Every 334,000 handgun purchase background checks in a state is associated with 86 suicides in that state per year, or every 3,900 background checks is associated with one suicide every year.

**TABLE 7**

**Fixed Effects Regression Results with Control Variables and Yearly Fixed Effects**

| | β | *t* | *p* |
|---|---|---|---|
| **Google Firearm Terms** | -.001 | -0.09 | 0.929 |
| | (-0.026 - 0.024) | | |
| **Divorce** | -.318 | -1.25 | 0.219 |
| | (-0.704 - 0.067) | | |
| **Unemployment** | .076 | 1.27 | 0.211 |
| | (-0.045 - 0.198) | | |
| **Binge** | .046 | 1.03 | 0.308 |
| | (-0.044 - 0.135) | | |
| **Mental State** | .051 | 0.20 | 0.845 |
| | (-0.467 - 0.568) | | |
| **Handgun Ownrshp** | **.029**** | 2.22 | 0.031 |
| | (0.003 - 0.056) | | |
| **BRFSS Dummy** | **2.49***** | 6.26 | 0.000 |
| | (1.689 - 3.285) | | |

Results are with state and year fixed effects. The dependent variable is the suicide rate per 100,000 population. Robust confidence intervals in parentheses. *** indicates significant at the 1% level. ** indicates significant at the 5% level. * indicates significant at the 10% level.

Lastly, the firearm-related terms show no association with the suicide rate however, handgun ownership is still found to be significantly associated with the suicide rate. We fail to reject the null hypothesis that firearm-related search terms are associated with the suicide rate. The firearm ownership proxy is again significant producing a slightly larger coefficient of 0.029. Thus, for every 334,000 handgun purchase background checks in a state is associated with 93 suicides in that state per year or every 3,600 background checks is associated with one suicide per year.

It is notable that the coefficients of both suicide and depression terms hardly change in models one and two, that is, the coefficients of the key predictor variables do not change when additional control variables are added. Thus, it appears that the influence of both suicide and depression terms are not confounded by any of the control variables.

## ROBUSTNESS CHECKS

It is possible that the search terms may be associated with a general aspect of an area rather than a specific phenomenon related to suicide. To test this possibility, we examined other possible health outcomes, cancer and overdose mortality, as dependent variables. If the search terms are significantly associated with such outcomes, then this paper cannot argue on the reliability of Google data and its association with the suicide rate.

Cancer mortality data was obtained from the CDC WONDER database. Cancer mortality is defined by ICD-10 codes C00-C97. Suicide-term and depression-term factors were tested individually in a bivariate fixed effects model controlling for state and yearly fixed effects. Suicide terms did not have an association with cancer mortality ($t$=0.70). Depression terms also did not have an association with cancer mortality ($t$=0.56).

It is conceivable that suicide and depression terms may be associated with similar outcomes such as drug overdoses. Given the rise in deaths of despair in the U.S. consisting of suicides, drug and alcohol poisonings, and liver diseases (Case and Deaton 2015), it is imperative that the variable of interest in this analysis only captures suicides, and does not capture other outcomes whose etiologies have been closely associated. All-intention drug overdose mortality data was obtained from the CDC WONDER database. All-intention drug overdose deaths are defined by ICD-10 codes X40-X44, X60-X64, X85, and Y10-Y14. Suicide-term and depression-term factors were tested individually in a bivariate fixed effects model controlling for state and yearly fixed effects. Suicide-terms and drug overdose deaths were not associated ($t$=-1.10). Depression-terms and drug overdose deaths were also not associated ($t$=-0.19). Thus, these analyses strengthen the assertion that search term data reveals profound social currents that dispose individuals, in this case, to the most personal of all choices, and that such associations are not spurious in nature.

## DISCUSSION

The first null hypothesis on suicide-terms can be rejected. The second null hypothesis on depression-terms can be rejected, however at the ten percent level. Lastly, we fail to reject the null of the third hypothesis on firearm-terms.

It is very well possible that collective sentiments within a state may guide individuals in that state to Google about suicide and suicide-related content. The extent to which Googling about suicide causes suicide and suicide attempts has not been fully determined. This analysis only provides a broad and collective lens into the suicidal inclinations of a state within a one year period. In other words, given that unemployment and divorce trends in a state takes the course of several months to years, the yearly time structure of this analysis may be capturing arregated collective sentiments that are associated with suicide. Given the impulsive nature of suicide (Deisenhammer et al. 2009), it can be inferred that if one does not have immediate access to a firearm, suicides may occur not long after seeking information online. Therefore, given that Google Trends data is available in real time and in shorter time frames, further research should have monthly and weekly time structures to capture more immediate sentiments that may shape the rate of suicide.

The fact that depression terms are slightly associated with the suicide rate give credence to the mixed literature that mental health state contributes to suicide. However, not all who are depressed experience suicide ideation and not all who experience suicide ideation may be depressed. The association between mental health state and suicide is complex and difficult to untangle at the present. As this analysis makes use of a digital window into social behavior, this analysis further confirms the complexities of this association.

Firearm terms were not associated with suicide rate which is a surprise given that half of all suicides are conducted via firearm and that the handgun ownership proxy is significant. One possible explanation for this may be that firearms and ammunition are more likely to be purchased at shops and not online. Also, firearms are enduring possessions in which one's purchase will provide an individual with the most effective means of suicide for decades to come. Therefore, after one possesses a firearm, they may leave very little digital traces afterwards. Lastly, firearm owners may tend to be male, older, and thus have less access to the internet.

As such, further research is needed to determine which age and gender populations are being captured most by Google search terms. Internet users tend to be younger although no significant gender differences on internet usage are reported. It is known that divorces and unemployment affect middle-aged men at disproportionately higher rates (Fincham 2011). This is because at middle-age, family and occupation constitute a larger part of one's identity, whereas younger individuals may find it easier to move into a new field or remarry. Therefore, anomic events such as a divorce or a layoff, will have more deleterious effects on middle-aged individuals than other age groups. Further research should parse out which age and gender groups Google search terms are most effective at measuring and which ones are they failing to measure.

This analysis demonstrates that online search behaviors go above and beyond the traditional measures used in suicidology. It can be inferred that the frequency of online searches can provide a lens to the collective thoughts, worries, and sentiments of the members of a society. That is, digital traces of the collective sentiments of an area can be analyzed for health outcome research. In this instance, the frequency with which suicidal search terms are queried in a state are significantly associated with the suicide rate of the state, even after controlling for unemployment, divorce, mental health state, binge drinking behaviors, handgun ownership, and state and yearly trends. Thus, this analysis identifies suicidal search term data as a potential leading measure of suicidal behavior. Search term data is more viable and more closely associated with the suicide rate than all other established measures.

Furthermore, search term data provides researchers with a more cost-effective method for assessing geographical sentiment, which is tied to health outcomes. The use of online digital methodologies should appeal to a wide range of researchers in the fields of public health, epidemiology, sociology, and psychology. In fact, user-posted content from social media site Twitter has allowed researchers to predict binge drinking behaviors (Curtis el al. 2018), onset of mental illness (Reece et al. 2017), and assessing cases of HIV (Young et al. 2014).

Specifically, search term data can be used to inform public health initiatives aimed at reducing suicide and assessing the prevalence of mental illness. Given the impulsive nature of suicide, firearm owners are more likely to use their firearms on themselves. However, non-firearm owning individuals who are suicidal may be more likely to search for suicide methods online. The time in which one gathers information about suicide and then commits the act is a window of opportunity for suicide prevention that digital methodologies provides. The fact that search term data is provided by Google in real-time makes this assertion more possible.

While strategies to reduce firearm suicides have been clearly spelled out by Kposowa et al. (2016) demonstrating that firearm availability and regulation have a significant impact on suicide rates, strategies on reducing non-firearm suicide rates may be informed by Google search term data. Non-firearm suicides include hanging, jumping, poisoning, cutting, overdosing, and suffocation, all means by which a successful completion requires some knowledge, and what better means to attain such knowledge than the most accessible library in existence, the internet.

In addition to demonstrating the reliability of Google search data as a measure of suicidality, the under-acknowledged firearm ownership proxy derived from firearm background checks reported by the FBI is also shown to be reliable. A preliminary analysis demonstrates this variable to be even more closely associated to the firearm-suicide rate. Thus, this analysis contributes to the public health literature by asserting that different methods of suicide call for different methods of prevention. This paper has touched upon two potential indicators of the suicide rate that complement each other.

Further preliminary analysis demonstrates that suicidal Google search terms have a much closer association to the non-firearm suicide rate and that the firearm ownership proxy has no association to the non-firearm suicide rate. These preliminary findings make sense given that individuals who do not own firearms are likely to search for effective suicide methods online. Likewise, individuals who do own firearms will most likely not Google search about suicide, especially when they are a click away from ending their lives. Such preliminary findings call for further research into which variables are more able to assess and predict suicide rates by method of completion.

Lastly, epidemiological and demographic contributions of this research may further inform researchers on where to target mental health and suicide prevention interventions. Areas that frequently query depressive and suicide-related content can be assessed in real time thus, improving current methods that rely on previous data. Search term data may provide insights into which search terms are the most appropriate to trigger a suicide-prevention screener such that an individual searching for suicide methods receives information on where to obtain help.

## LIMITATIONS

Given the yearly structure of this analysis, it ignores the effects that the variables may have on a shorter time scale and places arbitrary cutoff points in the longitudinal nature of the data. To address this limitation, further research should apply a monthly or quarterly structure to the data. Doing so would further elucidate the timing certain risk factors have on suicide rates. This assertion is supported by the literature suggesting the impulsive nature of suicide. The decision to commit suicide takes a shorter course than what is captured in a yearly analysis. This research should spur analyses implementing weekly or monthly time scales to the Google search term data.

The classic Durkheimian variable of religiosity (increases social integration leading to less suicide) was not used in the analysis as a result of a lack of time-variant data. The religious census does not provide enough years of data. The GSS is bi-annual, however it lacks enough statistical power to be representative at the state level. By omitting religiosity, it gets absorbed to some extent by the state fixed effects and holds the assumption that it is time-invariant, which is subject to much debate. Religiosity, measured by church attendance, has been declining for decades, and a noted increase in affiliation with no religion makes declining religiosity worthy of discussion. Furthermore, this study did not control for long term unemployment, which has been found to be associated with suicide (Kposowa et al. 2016)

Other limitations of this study include its small sample size of fifty states over eleven years (N: 550) and not accounting for smaller geographical aggregates such as county and media market areas. Future research should focus on these areas as many states have varying cultural climates within them and thus produce separate, yet consistent suicide rates. The factor analysis from this study was conducted on one year (2016) albeit using forty terms. A more through factor analysis consisting of more years would strengthen the methods. Finally, future research should look into the predictive power of the Google search term data.

## CONCLUSION

This analysis establishes two leading measures of the state suicide rate. Google search data goes above and beyond the established sociological, psychiatric, and epidemiological measures of state-level suicide. This analysis further legitimates the FBI's firearm background check system as a viable and applicable proxy for firearm ownership in efforts to measure state-level rates of suicide. Other attempts to assess firearm ownership in the U.S. have been deficient, thus this analysis further legitimates a new and more robust proxy for this elusive variable. As advances in technology become more utilized and institutionalized in society, researchers ought to make use of such technologies as they provide a deeper and more profound glimpse into everyday society.

*References*


Anon. 2018a. "Demographics of Mobile Device Ownership and Adoption in the United States." *Pew Research Center: Internet, Science & Tech*. Retrieved December 10, 2018 (http://www.pewinternet.org/fact-sheet/mobile/).

Anon. 2018b. "Demographics of Mobile Device Ownership and Adoption in the United States." *Pew Research Center: Internet, Science & Tech*. Retrieved December 3, 2018 (http://www.pewinternet.org/fact-sheet/mobile/).

Anon. 2018c. "Products - Data Briefs - Number 330 - September 2018." Retrieved December 3, 2018 (https://www.cdc.gov/nchs/products/databriefs/db330.htm).

Anon. n.d. "Google Market Share Countries 2018 | Statistic." *Statista*. Retrieved December 3, 2018 (https://www.statista.com/statistics/220534/googles-share-of-search-market-in-selected-countries/).

Arora, V. S., D. Stuckler, and M. McKee. 2016. "Tracking Search Engine Queries for Suicide in the United Kingdom, 2004-2013." *Public Health* 137:147–53.

Ayers, John W., Benjamin M. Althouse, Jon-Patrick Allem, J. Niels Rosenquist, and Daniel E. Ford. 2013. "Seasonality in Seeking Mental Health Information on Google." *American Journal of Preventive Medicine* 44(5):520–25.

Azrael, Deborah, Philip J. Cook, and Matthew Miller. 2004. "State and Local Prevalence of Firearms Ownership Measurement, Structure, and Trends." *Journal of Quantitative Criminology* 20(1):43–62.

Biddle, Lucy, Jane Derges, Becky Mars, Jon Heron, Jenny L. Donovan, John Potokar, Martyn Piper, Clare Wyllie, and David Gunnell. 2016. "Suicide and the Internet: Changes in the Accessibility of Suicide-Related Information between 2007 and 2014." *Journal of Affective Disorders* 190:370–75.

Case, Anne and Angus Deaton. 2015. "Rising Morbidity and Mortality in Midlife among White Non-Hispanic Americans in the 21st Century." *Proceedings of the National Academy of Sciences of the United States of America* 112(49):15078–83.

Chae, David H., Sean Clouston, Mark L. Hatzenbuehler, Michael R. Kramer, Hannah L. F. Cooper, Sacoby M. Wilson, Seth I. Stephens-Davidowitz, Robert S. Gold, and Bruce G. Link. 2015. "Association between an Internet-Based Measure of Area Racism and Black Mortality." *PloS One* 10(4):e0122963.

Chang, Shu-Sen, Simon Sai Man Kwok, Qijin Cheng, Paul S. F. Yip, and Ying-Yeh Chen. 2015. "The Association of Trends in Charcoal-Burning Suicide with Google Search and Newspaper Reporting in Taiwan: A Time Series Analysis." *Social Psychiatry and Psychiatric Epidemiology* 50(9):1451–61.

Choi, Hyunyoung and Hal Varian. 2009. "Predicting Initial Claims for Unemployment Benefits." *Google Inc* 1–5.

Choi, Hyunyoung and Hal Varian. 2012. "Predicting the Present with Google Trends." *The Economic Record* 88:2–9.

Cook, Philip J. and Jens Ludwig. 2006. "The Social Costs of Gun Ownership." *Journal of Public Economics* 90(1):379–91.

Curtis, Brenda, Salvatore Giorgi, Anneke E. K. Buffone, Lyle H. Ungar, Robert D. Ashford, Jessie Hemmons, Dan Summers, Casey Hamilton, and H. Andrew Schwartz. 2018. "Can Twitter Be Used to Predict County Excessive Alcohol Consumption Rates?" *PloS One* 13(4):e0194290.

Deisenhammer, Eberhard A., Chy-Meng Ing, Robert Strauss, Georg Kemmler, Hartmann Hinterhuber, and Elisabeth M. Weiss. 2009. "The Duration of the Suicidal Process: How Much Time Is Left for Intervention between Consideration and Accomplishment of a Suicide Attempt?" *The Journal of Clinical Psychiatry* 70(1):19–24.

Duggan, Mark. 2003. "Guns and Suicide." *Evaluating Gun Policy: Effects on Crime and Violence* 41–73.

Durkheim, Emile. 1951. "Suicide: A Study in Sociology (JA Spaulding & G. Simpson, Trans.)." *Glencoe, IL: Free Press. (Original Work Published 1897)*.

Fincham, Benjamin. 2011. *Understanding Suicide : A Sociological Autopsy*. Houndmills, Basingstoke, Hampshire; New York: Palgrave Macmillan.

Fox, Susannah and Harrison Rainie. 2014. *The Web at 25 in the US*. Pew Research Center [Internet & American Life Project].

Ginsberg, Jeremy, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. 2009. "Detecting Influenza Epidemics Using Search Engine Query Data." *Nature* 457(7232):1012–14.

Hagihara, Akihito, Shogo Miyazaki, and Takeru Abe. 2012. "Internet Suicide Searches and the Incidence of Suicide in Young People in Japan." *European Archives of Psychiatry and Clinical Neuroscience* 262(1):39–46.

Henry, Andrew F. and James F. Short. 1954. *Suicide and Homicide: Some Economic, Sociological and Psychological Aspects of Aggression*. Vol. 91442. Free Press New York.

Jun, Seung-Pyo, Hyoung Sun Yoo, and San Choi. 2018. "Ten Years of Research Change Using Google Trends: From the Perspective of Big Data Utilizations and Applications." *Technological Forecasting and Social Change* 130:69–87.

Kleck, Gary. 2004. "Measures of Gun Ownership Levels for Macro-Level Crime and Violence Research." *The Journal of Research in Crime and Delinquency* 41(1):3–36.

Knopov, Anita, Rebecca J. Sherman, Julia R. Raifman, Elysia Larson, and Michael B. Siegel. 2019. "Household Gun Ownership and Youth Suicide Rates at the State Level, 2005–2015." *American Journal of Preventive Medicine*.

Kposowa, Augustine, David Hamilton, and Katy Wang. 2016. "Impact of Firearm Availability and Gun Regulation on State Suicide Rates." *Suicide & Life-Threatening Behavior* 46(6):678–96.

Krumpal, Ivar. 2013. "Determinants of Social Desirability Bias in Sensitive Surveys: A Literature Review." *Quality & Quantity* 47(4):2025–47.

Lang, Matthew. 2013. “Firearm Background Checks and Suicide.” *The Economic Journal* 123(573):1085–99.

Lester, David and J. F. Gunn. 2016. “Sociology of Suicide.” Pp. 273–76 in *Stress: Concepts, Cognition, Emotion, and Behavior*. Elsevier.

Liu, Nancy H., Omar Contreras, Ricardo F. Muñoz, and Yan Leykin. 2014. “Assessing Suicide Attempts and Depression among Chinese Speakers over the Internet.” *Crisis* 35(5):322–29.

Mann, J. John and Christina A. Michel. 2016. “Prevention of Firearm Suicide in the United States: What Works and What Is Possible.” *The American Journal of Psychiatry* 173(10):969–79.

Matsubayashi, Tetsuya and Michiko Ueda. 2011. “The Effect of National Suicide Prevention Programs on Suicide Rates in 21 OECD Nations.” *Social Science & Medicine* 73(9):1395–1400.

Miller, Matthew, Deborah Azrael, and Catherine Barber. 2012. “Suicide Mortality in the United States: The Importance of Attending to Method in Understanding Population-Level Disparities in the Burden of Suicide.” *Annual Review of Public Health* 33:393–408.

Miller, Matthew, Deborah Azrael, and David Hemenway. 2002. “Household Firearm Ownership and Suicide Rates in the United States.” *Epidemiology* 13(5):517–24.

Miller, Matthew, Catherine Barber, Richard A. White, and Deborah Azrael. 2013. “Firearms and Suicide in the United States: Is Risk Independent of Underlying Suicidal Behavior?” *American Journal of Epidemiology* 178(6):946–55.

Miller, Matthew, Steven J. Lippmann, Deborah Azrael, and David Hemenway. 2007. “Household Firearm Ownership and Rates of Suicide across the 50 United States.” *The Journal of Trauma* 62(4):1029–34; discussion 1034–35.

Mok, Katherine, Anthony F. Jorm, and Jane Pirkis. 2015. “Suicide-Related Internet Use: A Review.” *The Australian and New Zealand Journal of Psychiatry* 49(8):697–705.

Mok, Katherine, Anna M. Ross, Anthony F. Jorm, and Jane Pirkis. 2016. “An Analysis of the Content and Availability of Information on Suicide Methods Online.” *Journal of Consumer Health on the Internet* 20(1-2):41–51.

Moksony, Ferenc and Rita Hegedűs. 2014. “The Use of Poisson Regression in the Sociological Study of Suicide.” *Corvinus Journal of Sociology and Social Policy* 5(2):97–114.

Neumayer, Eric. 2003. “Are Socioeconomic Factors Valid Determinants of Suicide? Controlling for National Cultures of Suicide with Fixed-Effects Estimation.” *Cross-Cultural Research: Official Journal of the Society for Cross-Cultural Research / Sponsored by the Human Relations Area Files, Inc* 37(3):307–29.

Norström, Thor. 1995. “The Impact of Alcohol, Divorce, and Unemployment on Suicide: A Multilevel Analysis.” *Social Forces; a Scientific Medium of Social Study and Interpretation* 74(1):293–314.

Paul, Elise, Roland Mergl, and Ulrich Hegerl. 2017. “Has Information on Suicide Methods Provided via the Internet Negatively Impacted Suicide Rates?” *PloS One* 12(12):e0190136.

Phillips, Michael Robert, Qijie Shen, Xiehe Liu, Sonya Pritzker, David Streiner, Ken Conner, and Gonghuan Yang. 2007. “Assessing Depressive Symptoms in Persons Who Die of Suicide in Mainland China.” *Journal of Affective Disorders* 98(1-2):73–82.

Recupero, Patricia R., Samara E. Harms, and Jeffrey M. Noble. 2008. “Googling Suicide: Surfing for Suicide Information on the Internet.” *The Journal of Clinical Psychiatry* 69(6):878–88.

Reece, Andrew G., Andrew J. Reagan, Katharina L. M. Lix, Peter Sheridan Dodds, Christopher M. Danforth, and Ellen J. Langer. 2017. "Forecasting the Onset and Course of Mental Illness with Twitter Data." *Scientific Reports* 7(1):13006.

Rice, Ronald E. 2006. "Influences, Usage, and Outcomes of Internet Health Information Searching: Multivariate Results from the Pew Surveys." *International Journal of Medical Informatics* 75(1):8–28.

Song, Tae Min, Juyoung Song, Ji-Young An, Laura L. Hayman, and Jong-Min Woo. 2014. "Psychological and Social Factors Affecting Internet Searches on Suicide in Korea: A Big Data Analysis of Google Search Trends." *Yonsei Medical Journal* 55(1):254–63.

Stephens-Davidowitz, Seth. 2014. "The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data." *Journal of Public Economics* 118:26–40.

Stephens-Davidowitz, Seth and Steven Pinker. 2017. *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us about Who We Really Are*. HarperCollins New York.

Sueki, Hajime. 2011. "Does the Volume of Internet Searches Using Suicide-Related Search Terms Influence the Suicide Death Rate: Data from 2004 to 2009 in Japan." *Psychiatry and Clinical Neurosciences* 65(4):392–94.

Sueki, Hajime. 2012. "Association between Deliberate Self-Harm-Related Internet Searches and the Mental States and Lifetime Suicidal Behaviors of Japanese Young Adults." *Psychiatry and Clinical Neurosciences* 66(5):451–53.

Sullivan, Danny. 2016. "Google Now Handles at Least 2 Trillion Searches per Year." *Search Engine Land* 24.

Thornton, Louise, Tonelle Handley, Frances Kay-Lambkin, and Amanda Baker. 2017. "Is A Person Thinking About Suicide Likely to Find Help on the Internet? An Evaluation of Google Search Results." *Suicide & Life-Threatening Behavior* 47(1):48–53.

Vitt, David C., Alexander F. McQuoid, Charles Moore, and Stephen Sawyer. 2018. “Trigger Warning: The Causal Impact of Gun Ownership on Suicide.” *Applied Economics* 50(53):5747–65.

Wasserman, I. M. 1984. “The Influence of Economic Business Cycles on United States Suicide Rates.” *Suicide & Life-Threatening Behavior* 14(3):143–56.

Wiebe, Douglas J. 2003. “Homicide and Suicide Risks Associated with Firearms in the Home: A National Case-Control Study.” *Annals of Emergency Medicine* 41(6):771–82.

Wintemute, G. J., S. P. Teret, J. F. Kraus, and M. W. Wright. 1988. “The Choice of Weapons in Firearm Suicides.” *American Journal of Public Health* 78(7):824–26.

Wray, Matt, Cynthia Colen, and Bernice Pescosolido. 2011. “The Sociology of Suicide.” *Annual Review of Sociology* 37(1):505–28.

Young, Sean D., Caitlin Rivers, and Bryan Lewis. 2014. “Methods of Using Real-Time Social Media Technologies for Detection and Remote Monitoring of HIV Outcomes.” *Preventive Medicine* 63:112–15.