Review of Economic Analysis 9 (2017) 107-125 1973-3909/2017107

Business 'Psych'cles: A Close Look at Mental Health and State-level Economic Performance Using Google Search Data

DANIEL FARHAT
Radford University *

TARJA VIITANEN
University of Otago *

This paper looks at the mental health of US consumers and the state-level economy between 2004 and 2014. To better capture the mental condition of Americans, we use internet query data for select psychological keywords. We evaluate the seasonal fluctuations in this data and compare the broader trends to indices of economic misery. Our results reaffirm a seasonal period of despair from late autumn to early spring. Further, we show that economic misfortune is positively related to online interest in psychological health. The results provide valuable input to public health providers for forming 'nowcasts' of mental health needs and to health policy-makers for forming spatially- and temporally-targeted mental health initiatives.

Keywords*:* mental health, business cycles, internet keyword search data

JEL Classifications*:* I18, J11, C81

1 Introduction

Sadness is expensive. Mental and substance-use disorders account for approximately 8% of the global burden of all disease, affecting as many as 700 million people worldwide. Depression is one of the leading causes of disability worldwide[1], leading to lost production. Recent estimates rank mental and substance-use disorders to be third in the leading global causes of disability-adjusted life years, accounting for 23% of all years lived with disability (Ferrari *et al.*, 2014; Whiteford *et al.*, 2013). The economic burden imposed on the global society amounted to an

* We would like to thank the University of Otago UORG grant for financial assistance. We have no additional competing interests.

[1] http://www.economistinsights.com/healthcare/event/global-crisis-depression

estimated US$2.5 trillion in 2010 and are forecast to reach US$6.0 trillion by 2030 (Bloom *et al*., 2011). These figures may underestimate the true burden of poor mental health as psychological problems are difficult to diagnose, and many opt to suffer in silence (Prince *et al*., 2007; Patel *et al.*, 2014).

Waves of sadness predictably hit Americans each year. The seasonal variation in mental health is known amongst the general public as "Seasonal Affective Disorder" (SAD). These seasons coincide with the late autumn, winter, and early spring months when the weather is generally gloomy. Other factors contribute to particularly high or low mental states across years. One such factor is the stage of the business cycle (see Zivin *et al.* (2011) for a review of the previous literature). People are generally happier during periods of economic prosperity and sadder during periods of economic decline.

Many of the observed trends in mental health are based on diagnoses or surveys. As a result, people who are experiencing symptoms of poor mental health but have not sought treatment or have not been diagnosed are not accounted for. The stigma associated with mental health problems discourages people from discussing their feelings with medical professionals. The advent of internet technology, however, has both increased people's access to psychology-related information for private self-diagnoses (Tefft, 2011) *and* given researchers a way of measuring general interest in any topic (by recording the frequency of keywords used on search engines).

This paper proposes a novel way to target interventions at the regional level across time. We use Google search data for select keywords as an indicator of latent demand for mental health care services. We evaluate the trends in these searches across US states and compare them to broad movements in the states' economies. The results can be used to inform policy. Katitireddi *et al.* (2012) urges for monitoring of recessionary impacts on health inequalities in response to ongoing labour market and social policy changes resulting from the Great Recession. Using real-time Google search data enables policy makers to provide the most targeted help to those in need at a potentially low cost, for example, by using targeted internet advertising regarding help/hotlines.

2 Previous Literature

2.1 Seasonal Variation in Mental Health Disorders

Seasonal depression accounts for approximately 11% of all cases of major depressive episodes and affects women relatively more than men (Levitt *et al.*, 2000). The prevalence of SAD in the general population is significant, ranging from estimated 2.4%-2.6% in North Wales and Canada respectively (Michalak *et al.*, 2001, Levitt and Boyle, 2002) to about 12% among rural Finns and Lapps (Saarijarvi *et al.*, 1999). It appears to be largely underdiagnosed and/or misdiagnosed (Michalak *et al.*, 2001).

There is a spatial element to SAD. Mersch *et al.* (1999) reviews the earliest literature on the impact of latitude on SAD and summarises that the mean prevalence of SAD is twice as high in North America (in the west) than in Europe (in the east). Overall, the correlation between prevalence and latitude (north to south) was not found to be significant in the reviewed studies. However, a small significant positive correlation was found between prevalence and latitude within North America with the trend in Europe heading in the same direction. In a more recent, randomised phone survey conducted in Ontario, Canada, Levitt and Boyle (2002) discovered that the prevalence of seasonal depression was not influenced by latitude in their sample; however, since sampling occurred only over 8 degrees of latitude it is possible that there may be a significant effect of latitude on SAD over a greater range of latitudes.

Overall, SAD is found to be correlated with a wide variety of other factors. Female gender, high body mass index, high level of education and young age are all positively and significantly related to SAD (Saarijarvi *et al.*, 1999). Also it is also suggested that overall climate, genetic vulnerability and socio–cultural context may play a role as a determinant of SAD (Mersch *et al.*, 1999).

An extreme outcome of depression is suicide. Several studies have found a seasonal component in suicide occurrence. Australia, which consistently ranks as one of the happiest and most satisfied countries in the world (Leigh and Wolfers, 2006), has a dark underside of increasing seasonal rates of suicide (Rock *et al.*, 2003). Rock *et al.* (2003) find a peak in suicides during spring time (September-November in Australia) with this effect produced by men who use impulsive, violent methods. Similarly an earlier study from the 1970s for Australia found a spring effect but only for females (Parker and Walter, 1982). A spring peak in suicides has been observed as early as the 19th century in Finland, with suicide rates amongst the highest in the world with about 70% violent suicides (Rasanen, 2002). Rasanen *et al.* (2002) further confirms the seasonality of suicides in Finland (with a peak from April to July for men and peaks in May and October for women), however that there is a significant summer peak for males who commit suicide by gas poisoning (a non-violent method). Additional factors contributing to the seasonality of suicides include differences in patterns of anti-depressant use and immigration patterns (Rock *et al.*, 2003).

2.2 Business Cycle Variation in Mental Health Disorders

The impact of business cycles on social conditions such as crime, mortality and divorce has been of interest to statisticians since the 1920s (Ogburn and Thomas, 1922; Thomas, 2015). The underlying mechanism is explained by Brenner (1979), who posits that during an economic recession individuals with less economic security experience more stress. He further explains how this stress can lead to social and family structure breakdown as well as an adoption of habits which are harmful to health; a combination of all of the above can lead to an extreme event such as a suicide. Unintentional deaths, like drunk-driving fatalities, also increase. Ruhm (1995), for example, finds that alcohol consumption and traffic deaths vary pro-cyclically.

In fact, depression and suicide are found to depend on the business cycle in the US with working age individuals' (age 25-64) rate of suicide falling during economic expansions and rising in recessions (Luo *et al.*, 2011). Furthermore, Yang (1992) finds that unemployment rate has a significant detrimental impact on white male suicide rate whereas female labour force participation rate reduced female suicide rates.

The Global Financial Crisis (GFC) and its effect on mental health has been of interest to several researchers, including Katikireddi *et al.* (2012), Lee *et al.* (2010), Gili *et al.* (2013), Economou *et al.* (2013), and Sargent-Cox *et al.* (2011). Katikireddi *et al.* (2012) find that British men's mental health regardless of their employment status has deteriorated within two years of the onset of the GFC. Similarly, in Sargent-Cox *et al.* (2011) the GFC increased depressive and anxiety symptoms in a sample of older Australians. Spain has been one of the worst hit after the GFC. Gili *et al.* (2013) find that the recession has increased the frequency of mental health disorders and alcohol abuse in Spain, especially among families experiencing unemployment and mortgage payment difficulties.

Further, Lee *et al.* (2010) find that the Global Financial Crisis has been associated with a significant increase in the risk for depression in Hong Kong especially for the employed, home-makers, high to middle income and married individuals. Similarly for Greece, Economou *et al.* (2013) find that the prevalence rate for major depression increased from 3.3% in 2008 to 8.2% in 2011 with largest effects for young people, married individuals, those with financial distress and people who use medication.

2.3 The Use of Internet Data for Self-Help

The internet has dramatically increased peoples' access to information. We can meter how often people access data online using metrics on the relative frequency of keyword searches on search engines (like Google). Choi and Varian (2012) demonstrate how Google keyword search data can be used for contemporaneous forecasting (or 'nowcasting') of economic indicators such as automobile sales, unemployment claims and consumer confidence. Models that include Google keyword search data outperform those that exclude these predictors by 5-20%. This data has also been used to predict spreading of seasonal influenza (Ginsberg *et al.*, 2009) and financial market fluctuations (Preis *et al.*, 2010).

In particular, the internet has dramatically reduced the costs of accessing information regarding health. In 2001 approximately 40% of surveyed individuals with internet access reported using the internet to look for advice or information on health or health care (Baker *et al.* 2003) while by 2008 it has risen to 61% (Fox and Jones, 2009). Ginsberg *et al.* (2009) state that the relative frequency of certain Google search queries is highly correlated with the percentage of physician visits where the patient presents with influenza-like symptoms indicating a complementary nature of health-related internet searches and demand for health care (Suziedelyte, 2012).

McMullan (2006) shows that the majority of health-related searches are for specific conditions, and suggests that people turn to the internet first for two reasons: 1) an attempt to manage their own health care independently and/or to decide whether they need professional help before a clinical encounter and 2) to find out more information on the condition after a clinical encounter. We might delve into the motivations further. One reason for preliminary self-diagnosis might be related to travel costs. Bundorf *et al.* (2006), for example, find that uninsured individuals with reported chronic health conditions and those with longer travel times for their usual source of health care were more likely to search for health information on the internet. The anonymous nature of internet-based self-diagnosis is also appealing. For stigmatised illnesses, patients may attempt to avoid embarrassment or shame by self-diagnosing online before discussing their symptoms unnecessarily with a medical professional. Certain groups in the population might be more likely to do this than others. For example, Gould *et al.* (2002) find that especially adolescents search for answers on the internet for emotional problems.

In a similar analysis to this study, Tefft (2011) examines the association between weekly unemployment insurance claims, monthly unemployment rates and Google search indices for depression and anxiety. Tefft (2011) justifies the use of Google search data as “meaningful representations of the intent to understand or seek treatment for symptoms of psychological distress that are experienced at the time of search” (p.258). Our paper expands the analysis of Tefft (2011) by allowing for the widely documented seasonal variation in mental health occurrence as well as allowing for a wider variety of search terms and a longer time period. Furthermore, whereas Tefft (2011) concentrates on unemployment related effects on depression and anxiety searches, we look at the state of the economy in general that will have effects on the larger population rather than just those who have lost their jobs or are at risk of losing their jobs. We also provide a detailed spatial analysis that has not been done in the previous literature.

3 Data Analytic Procedures

Google has been collecting data on the frequency (relative to all Google searches) of the keywords used on its search engine since 2004. Because these searches are passively monitored, this data gives us an unadulterated view into what is capturing people’s attention. Google makes this data freely available through a web-based retrieval tool called *Google Trends*. If you specify a keyword of interest, Google Trends will provide you with an index representing how often that term had been searched for relative to all Google searches. This data is available at the weekly or monthly frequency (depending on availability). Google Trends has global reach; data can be categorised by nation, state and (in some cases) city. The data is also categorised by topic (e.g. “Health” or “Shopping”) to allow us to refine the meaning behind the searches. See the appendix at the end of this article for additional details.

In this study, we focus on four particular keywords: “anxiety”, “depression”, “stress” and “suicide”. We extract a time-series from Google Trends for each of these keywords, for each US state. Our query is limited to the “Health” category. Weekly data is available for more populous states, while less populous states tend to only have monthly data. Weekly data is therefore averaged to form a monthly approximation so that all series have the same frequency. We adopt a standard normal transformation for each of the series in the panel in a way that the standardised distribution for each state across time has a mean zero and a variance of one. To visualise the level of interest in these terms across space and time, we construct a series of animated choropleth maps. These are available for viewing at https://sites.google.com/site/vfresearch2015/.

To measure overall economic performance, we use a coincident index provided by the Federal Reserve Bank of Philadelphia. To construct the index, four monthly economic indicators are combined using a dynamic single-factor model[2]: nonfarm payroll employment, average hours worked in manufacturing, the unemployment rate and salary disbursements (deflated by the CPI). The trend for each state’s index is set to match that of its gross domestic product. All states except District of Columbia are represented. The index is consistent with a positive view of the economy, meaning that the index number rises when the state economy performs well and falls when the state economy performs poorly. We take an inverse of the index to form a “misery index.” This data has been useful for assessing business cycle movements in individual states, which do not perfectly sync with each other or with national business cycles (Crone, 2006).

Summary statistics for our data set are reported in Table 1. As an illustration, we show the Google Trends normalised indices and the economic index for Texas in figure 1. No obvious correlation exists in Texas at first glance, perhaps because seasonal effects and a long-term trend in the search indices obscure the relationship.

We estimate the following model to assess the relationship between the misery index, seasons and a collection of Google search terms related to mental health issues:

$$Google\ search\ index_{imt} = Misery\ index_{imt} + \delta_i + \gamma_m + \theta_t + \varepsilon_{imt} \tag{1}$$

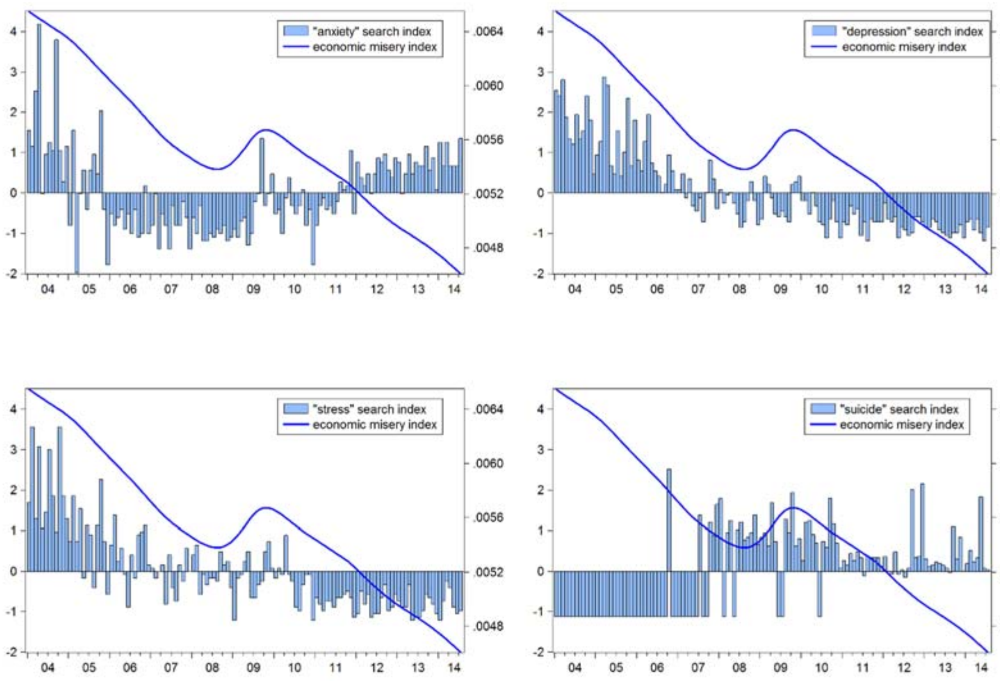
where Google search index is a normalised index for stress, anxiety, depression or suicide in a given state (i) for a given month (m) in a given year (t).

[2] For more information on the construction of these indices, see Crone and Clayton-Matthews (2005).

Table 1: Summary statistics

	N	Mean	Std. Dev.	Min	Max
Misery index	4459	-160.914	23.511	-259.950	-105.070
“Stress” search index	4459	0.000	0.995	-4.474	4.704
“Anxiety” search index	4459	0.000	0.995	-5.479	3.282
“Depression” search index	4459	0.000	0.995	-4.092	4.635
“Suicide” search index	3458	0.000	0.995	-2.522	9.435

Figure 1: Data from Texas on Internet Search
(“Anxiety”, “Depression”, “Stress” and “Suicide”) and Economic Activity (Misery):



The misery index is also provided by state, month, and year. Also, δ_i includes state fixed effects, γ_m controls for seasonal (month) effects relative to December and θ_t controls for annual effects.

4 Results and Discussion

The regression results below report the statistical relationship between the Google Trends indices for mental health keywords, the season, the location and the state of the economy. Our results suggest that statistically significant relationships exist on nearly all dimensions.

Table 2 shows the results for "stress" searches. The worsening of the local economy, as proxied by the state-level misery index, has a highly significant effect in individuals' Google searches for information on stress. There is also interesting seasonal variation in stress related Google searches with less searches during the summer months (June-August) and more intense stress-related searches in the winter (February to April) and again in the autumn (September to October).

Table 3 shows similar results for "anxiety" in such way that the misery index positively affects Google searches for anxiety with the effect being highly significant and positive. The seasonal variation indicates more anxiety-related Google searches throughout the year except for early summer (May-June) which fail to be significantly different from zero.

The results for "depression", which appear in Table 4, show a significant positive relationship between the misery index and Google searches. Again, as the economy worsens the frequency of "depression" searches rise. These searches also vary significantly by season with the most searches during the autumn months (October-November) and winter months (January-April) which is consistent with the previous literature on the Seasonal Affective Disorder.

In extreme cases, stress, anxiety and depression can lead to suicide. The results for "suicide" searches in appear in Table 5 indicate that suicides are either not affected by the local economy or that individuals do not use Google to search for "self-help" on suicide. Indeed, Zivin *et al.*(2011) discuss that suicide is "an extreme outcome and rare event" (p.1346). However, there is a seasonal peak in searches for suicide in September which may coincide with the start of a school/college year for many students.

Table 2: "Stress" Index Results: Seasonal and Economic Cycle Effects

	OLS Coefficient	**SE**
Misery index	0.010***	0.002
January	-0.090	0.066
February	0.303***	0.066
March	0.441***	0.066
April	0.653***	0.066
May	0.098	0.066
June	-0.326***	0.066
July	-0.384***	0.066
August	-0.293***	0.068
September	0.474***	0.068
October	0.693***	0.068
November	0.631***	0.068
N	4459	

Note: Regressions also include state fixed effects and year fixed effects.

Table 3: "Anxiety" Index Results: Seasonal and Economic Cycle Effects

	OLS Coefficient	**SE**
Misery index	0.009***	0.002
January	0.158***	0.057
February	0.239***	0.057
March	0.246***	0.057
April	0.260***	0.057
May	0.060	0.057
June	-0.016	0.057
July	0.127**	0.057
August	0.311***	0.058
September	0.420***	0.058
October	0.443***	0.058
November	0.378***	0.058
N	4459	

Note: Regressions also include state fixed effects and year fixed effects.

Table 4: "Depression" Index Results: Seasonal and Economic Cycle Effects

	OLS Coefficient	**SE**
Misery index	0.009***	0.002
January	0.148**	0.067
February	0.267***	0.067
March	0.313***	0.067
April	0.371***	0.067
May	-0.079	0.067
June	-0.499***	0.067
July	-0.382***	0.067
August	-0.460***	0.069
September	0.100	0.069
October	0.433***	0.069
November	0.517***	0.069
N	4459	

Note: Regressions also include state fixed effects and year fixed effects.

State-level seasonal variation is analysed in Tables 6-9. Table 6 for stress searches shows the coefficient and the standard error of the state dummy from the previous regressions. There appears to be significant positive effects on stress searches from residing in the Western states excluding California and Washington.

Table 5: "Suicide" Index Results: Seasonal and Economic Cycle Effects

	OLS Coefficient	SE
Misery index	0.003	0.003
January	-0.294***	0.080
February	-0.116	0.080
March	-0.075	0.080
April	0.026	0.080
May	-0.311***	0.080
June	-0.415***	0.080
July	-0.484***	0.080
August	-0.411***	0.082
September	0.284***	0.082
October	0.052	0.082
November	0.010	0.082
N	3458	

Note: Regressions also include state fixed effects and year fixed effects.

The negative/positive month effects in Tables 6-9 are based on a separate regression where the state is interacted with the month and where the negative/positive status is based on the sign of the coefficient when it is significant at 10% level or less. "Stress" keyword searches appear to reduce in the summer months and increase in the winter and spring time. This could be explained by the prevalence of SAD in the winter months, however, it does not appear to be that straightforward as also some of the sunnier states in the south follow the same pattern (namely New Mexico). "Anxiety" related searches are most common in March and September. "Depression"-related searches reduce in the summer months as expected, possibly due to the dual effect of more sunshine hours as well as holidays. "Suicide" searches seem most common in April; this result confirms the spring peak in suicides found in the previous literature.These results between Google searches for keywords of interest and the state of the local economy and season provide valuable input into public health providers with a potential to nowcast mental health needs and help policy makers target interventions by region, season and the economic cycle. Considering that many of the individuals who use internet searches as a form of "self-help" may never reach out to a professional health care professional it is crucial that the health care interventions can try to reach out to these people. An obvious mechanism for reaching these individuals is to use targeted internet advertising on, for example, anonymous help lines.

Table 6: Spatial Analysis of "Stress" Searches

	OLS Coef.	SE	Negative month effects	Positive month effects
Alaska	-0.122	0.135	September	
Arizona	0.624	0.193***		
Arkansas	0.104	0.134		
California	0.180	0.138		May
Colorado	0.381	0.157**		November
Connecticut	0.093	0.134	July	
Delaware	0.073	0.133	August	
Florida	0.174	0.138		
Georgia	0.233	0.142		January, March, November
Idaho	0.566	0.183***		
Illinois	0.012	0.132		March
Indiana	0.010	0.132	June	
Iowa	0.075	0.133		
Kansas	0.036	0.132		
Kentucky	-0.024	0.132		March
Louisiana	-0.019	0.132		
Maine	-0.034	0.132		March
Maryland	0.157	0.137	July	
Massachusetts	0.240	0.143*		
Michigan	-0.203	0.140		March, November
Minnesota	0.149	0.136	June, July, August	
Mississippi	0.036	0.132		
Missouri	-0.018	0.132		March, May, November
Montana	0.513	0.175***	August, September	
Nebraska	0.206	0.140		
Nevada	0.702	0.206***		
New Hampshire	0.413	0.161**		March, May
New Jersey	0.066	0.133	July, August	
New Mexico	0.288	0.147**		March, November
New York	0.031	0.132	August	May
North Carolina	0.215	0.141		
North Dakota	0.431	0.164***	September	
Ohio	-0.064	0.133		January, March, May
Oklahoma	0.175	0.138		March
Oregon	0.526	0.177***		January, March, May
Pennsylvania	0.015	0.132	June, July, August	
Rhode Island	0.087	0.133	July, September	
South Carolina	0.133	0.135		
South Dakota	0.325	0.151**		
Tennessee	0.117	0.135		January, March
Texas	0.465	0.168***		
Utah	0.545	0.180***		January, March
Vermont	0.113	0.134		January, March
Virginia	0.185	0.138		
Washington	0.111	0.134		March
West Virginia	0.025	0.132		March, May
Wisconsin	0.355	0.154**	June, July, August	
Wyoming	0.570	0.184***		

Notes: "Negative month effects" refers to the sign of a statistically significant coefficient between state and month and vice versa for "Positive month effects".

Table 7. Spatial Analysis of "Stress" Searches

	OLS Coef.	SE	Negative month effects	Positive month effects
Alaska	-0.119	0.115	August, October	
Arizona	0.607	0.165***	January	
Arkansas	0.101	0.115		
California	0.175	0.118		
Colorado	0.370	0.135***		
Connecticut	0.090	0.114		
Delaware	0.071	0.114		
Florida	0.169	0.118		June, September
Georgia	0.226	0.122*	February	
Idaho	0.551	0.157***		June, July
Illinois	0.011	0.113		
Indiana	0.010	0.113		
Iowa	0.073	0.114		
Kansas	0.035	0.113		September
Kentucky	-0.023	0.113		
Louisiana	-0.018	0.113		
Maine	-0.033	0.113		
Maryland	0.153	0.117		March
Massachusetts	0.234	0.122*		March
Michigan	-0.197	0.120*	May	
Minnesota	0.145	0.117	November	
Mississippi	0.035	0.113		
Missouri	-0.018	0.113		March
Montana	0.499	0.150***		
Nebraska	0.200	0.120*		March
Nevada	0.683	0.176***	October	
New Hampshire	0.402	0.138***		
New Jersey	0.064	0.114		September
New Mexico	0.280	0.126**		
New York	0.030	0.113		
North Carolina	0.209	0.120*		September
North Dakota	0.419	0.140***		
Ohio	-0.062	0.114		March, September
Oklahoma	0.170	0.118		
Oregon	0.512	0.152***		
Pennsylvania	0.015	0.113		September
Rhode Island	0.084	0.114		
South Carolina	0.129	0.116		
South Dakota	0.316	0.129**		
Tennessee	0.114	0.115		
Texas	0.452	0.144***		
Utah	0.530	0.154***		
Vermont	0.110	0.115		
Virginia	0.180	0.118		
Washington	0.108	0.115		
West Virginia	0.024	0.113		
Wisconsin	0.345	0.132***	June	
Wyoming	0.555	0.158***		

Table 8: Spatial Analysis of "Anxiety" Searches

	OLS Coef.	SE	Negative month effects	Positive month effects
Alaska	-0.12	0.136	March, April, July	
Arizona	0.612	0.195***		
Arkansas	0.102	0.136		
California	0.177	0.140		
Colorado	0.373	0.159**	June	
Connecticut	0.091	0.135	July	
Delaware	0.071	0.135		
Florida	0.170	0.139	April, July	
Georgia	0.228	0.144		
Idaho	0.555	0.186***		
Illinois	0.011	0.134		
Indiana	0.010	0.134		
Iowa	0.074	0.135		
Kansas	0.035	0.134	March, July, August	
Kentucky	-0.023	0.134	March, April, July	
Louisiana	-0.018	0.134		
Maine	-0.033	0.134		
Maryland	0.154	0.138	July	
Massachusetts	0.235	0.144		
Michigan	-0.199	0.141	July, August	
Minnesota	0.146	0.138	July	
Mississippi	0.035	0.134	July	
Missouri	-0.018	0.134		
Montana	0.503	0.177***		
Nebraska	0.202	0.141		
Nevada	0.689	0.208***		
New Hampshire	0.405	0.163**		
New Jersey	0.065	0.134	February, March, April, July, August, September	
New Mexico	0.283	0.149*	July, August	
New York	0.030	0.134		
North Carolina	0.211	0.142	July, August	
North Dakota	0.422	0.166*		
Ohio	-0.062	0.134	April, July, October	
Oklahoma	0.172	0.139	July	
Oregon	0.516	0.179***	January, July	
Pennsylvania	0.015	0.134		
Rhode Island	0.085	0.135		
South Carolina	0.130	0.137	July	
South Dakota	0.318	0.153**	March	
Tennessee	0.115	0.136	July	
Texas	0.456	0.170***		
Utah	0.535	0.182***		
Vermont	0.111	0.136	August, September	
Virginia	0.181	0.140		
Washington	0.109	0.136		
West Virginia	0.024	0.134	March, April, September	
Wisconsin	0.348	0.156**	March, April, September	
Wyoming	0.559	0.186***	March, April, July, August, September	

Table 9: Spatial Analysis of "Suicide" Searches

	OLS Coef.	SE	Negative month effects	Positive month effects
Arizona	0.191	0.248		
Colorado	0.117	0.188		
Connecticut	0.029	0.144		
Delaware	0.022	0.143		
Florida	0.053	0.152		
Georgia	0.071	0.160		
Idaho	0.174	0.233		April
Illinois	0.004	0.141		
Indiana	0.003	0.141		
Iowa	0.023	0.143		
Kansas	0.011	0.141		
Kentucky	-0.007	0.141		
Louisiana	-0.006	0.141		
Maryland	0.048	0.150		
Massachusetts	0.074	0.161		
Michigan	omitted			
Minnesota	0.046	0.149		
Mississippi	0.011	0.141		
Missouri	-0.006	0.141		
Nebraska	0.063	0.156		
Nevada	0.215	0.270		
New Hampshire	0.127	0.195		
New Jersey	0.020	0.142		
New Mexico	0.088	0.169		
New York	0.009	0.141		
North Carolina	0.066	0.157		
Ohio	-0.019	0.142		
Oklahoma	0.054	0.152		
Oregon	0.161	0.223		October
Pennsylvania	0.005	0.141		April
Rhode Island	0.027	0.143		
South Carolina	0.041	0.147		
Tennessee	0.036	0.146		
Utah	0.167	0.227		
Vermont	0.035	0.145		April
Virginia	0.057	0.153		April, May
Washington	0.034	0.145		April
West Virginia	0.008	0.141	September	
Wisconsin	omitted			

5 Implications and Conclusions

This paper has provided useful information to health care administrators interested in following the state of the nation's mental health to target interventions for those individuals who may not want to reach out to a mental health professional. We use publicly available data to examine how certain Google search keywords vary both by the economic cycle by state-level and by season. We find significant effects of the economic cycle on searches for stress, anxiety and depression. Using Google searches to help public health officials to target interventions is merely a start to what can be done. Data mining the internet could potentially help public officials to tackle issues such as school shootings as many of these offenders have reportedly discussed their intentions in online forums.

We also find interesting seasonal variation that can help public health officials target interventions for at-risk individuals. It is also important to note that once more data becomes available it is possible to target even smaller geographical areas to get most accurate information. Also, this paper relies on a relatively few years of monthly data, however, it may be possible in the future to use weekly or even daily data to efficiently target interventions. Mental health issues are a very serious concern for both developed and less-developed societies and due to the private nature of the perceived issues the method suggested in this paper can provide valuable input into tackling a global health care issue.

Appendix: Google Trends

Google Trends collects high frequency data (weekly or monthly). For any search term you ask for (X), Google Trends first finds the week/month with the highest search volume (let's call it X*) relative to all Google searches then indexes all of the other weeks relative to that week. Consider, for example, searches for the term "prescription drugs" in California (see Table A1).

Table A1: Data from California on the search term "prescription drugs":

WEEK	GOOGLE TRENDS INDEX
2004-03-07 - 2004-03-13	56
2004-03-14 - 2004-03-20	56
2004-03-21 - 2004-03-27	100
2004-03-28 - 2004-04-03	55
2004-04-04 - 2004-04-10	39
2004-04-11 - 2004-04-17	42
2004-04-18 - 2004-04-24	49

The week of March 21, 2004 had the highest number of searches for that term relative to all Google searches. Let's pretend that there were 5,000 total Google searches for all keywords and 350 were for "prescription drugs", so that week would have a score of 0.07 (or 7%). In the next week, the week of March 28th, there were 55% as many relative searches. This means that the week of March 28th should have a score of 0.55 x 0.07 = 0.0385 (= 3.85%). This could be either because there were fewer total searches for "prescription drugs" or because searches for "prescription drugs" didn't keep up with searches on Google in general (we do not know by looking at the number). As such, the Google Trends number represents particular searches in relation to 'overall' activity.

If you request data for more than a single keyword or more than a single location on Google Trends, all the data is rescaled to be relative to the most popular period/location in the time-span requested. For example, suppose you request weekly data for "depression" and "suicide" for California for 2009. If you look at the week of March 1-7, 2009, the "depression" measure would be 100 and the "suicide" measure would be 75. In the next week, these numbers are 85 and 76 respectively. These numbers must be interpreted cautiously. Google Trends has identified the most popular instance of the most popular keyword. In all of 2009, the most frequent search out of both terms was for "depression" during the first week of March. Google Trends has thus assigned this a value of 100 and has rescaled all the other "depression" data relative to the searches for "depression" in this week. Therefore, the score of 86 in the second week of March means that searches in this week were 85% of those in the first week of March. Not only is this done for the "depression" series, but it is also done for the "suicide" series. The measures of 75 and 76 for "suicide" during the first and second week of March means that searches for "suicide" were 75% and 76% of those for "depression" in the first week of March. The "suicide" sequence does not contain a 100 since suicide was not the most popular term in the time period requested. This re-scaling works in a similar fashion if you request more than a single location. Unfortunately, you can only request data for a limited number of keywords and locations at a time on Google Trends. For a large, comprehensive dataset, you would have to perform multiple requests, each of which would be rescaled separately. To prevent this re-scaling from introducing distortion in the data, data for each keyword and each location are requested one at a time. As a result, each keyword for each location is scaled relative to itself only.

References

Baker, L., Wagner, T. H., Singer, S., and Bundorf, M. K. (2003). Use of the Internet and E-mail for Health Care Information: Results from a National Survey, *Jama*, 289(18), 2400-2406.

Bloom, D. E., Cafiero, E. T., and Jané-Llopis, E. (2011). The Global Economic Burden of Non-Communicable Diseases, Geneva: World Economic Forum.

Brenner, M. H. (1979). Mortality and the National Economy: A Review, and the Experience of England and Wales, 1936-76, *The Lancet*, 314(8142), 568-573.

Bundorf, M. K., Wagner, T. H., Singer, S. J., and Baker, L. C. (2006). Who Searches the Internet for Health Information?, *Health Services Research*, 41(3p1), 819-836.

Choi, H., and Varian, H. (2012). Predicting the Present with Google Trends, *Economic Record*, 88(s1), 2-9.

Crone, T. M., and Clayton-Matthews, A. (2005). Consistent Economic Indexes for the 50 States, *Review of Economics and Statistics*, 87(4), 593-603.

Crone, T. M. (2006). What a New Set of Indexes Tells us About State and National Business Cycles, *Business Review*, Q1, 11-24.

Economou, M., Madianos, M., Peppou, L. E., Patelakis, A., and Stefanis, C. N. (2013). Major Depression in the Era of Economic Crisis: A Replication of a Cross-sectional Study Across Greece, *Journal of Affective Disorders*, 145(3), 308-314.

Federal Reserve Bank of Philadelphia (2015). *State Coincident Indices*, available at www.philadelphiafed.org.

Ferrari, A. J., Norman, R. E., Freedman, G., Baxter, A. J., Pirkis, J. E., Harris, M. G., et al. (2014). The Burden Attributable to Mental and Substance Use Disorders as Risk Factors for Suicide: Findings from the Global Burden of Disease Study 2010, *PLoS One*, 9(4): e91936.

Fox, S., and Jones, S. (2012). The Social Life of Health Information, *Pew Internet and American Life Project 2009*.

Gili, M., Roca, M., Basu, S., McKee, M., and Stuckler, D. (2013). The Mental Health Risks of Economic Crisis in Spain: Evidence from Primary Care Centres, 2006 and 2010, *The European Journal of Public Health*, 23(1), 103-108.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L. (2009). Detecting Influenza Epidemics Using Search Engine Query Data, *Nature*, 457(7232), 1012-1014.

Google (2014). *Google Trends*, available at www.google.com/trends.

Gould, M. S., Munfakh, J. L. H., Lubell, K., Kleinman, M., and Parker, S. (2002). Seeking Help from the Internet During Adolescence, *Journal of the American Academy of Child and Adolescent Psychiatry*, 41(10), 1182-1189.

Katikireddi, S.V., Niedzwiedz, C.L., and Popham, F. (2012). Trends in Population Mental Health Before and After the 2008 Recession: A Repeat Cross-sectional Analysis of the 1991–2010 Health Surveys of England, *BMJ Open*, 2(5): e001790.

Lee, S., Guo, W. J., Tsang, A., Mak, A. D., Wu, J., Ng, K. L., and Kwok, K. (2010). Evidence for the 2008 Economic Crisis Exacerbating Depression in Hong Kong, *Journal of Affective Disorders*, 126(1), 125-133.

Leigh, A., and Wolfers, J. (2006). Happiness and the Human Development Index: Australia is Not a Paradox, *Australian Economic Review*, 39(2), 176-184.

Levitt, A. J. and Boyle, M. H. (2002). The Impact of Latitude on the Prevalence of Seasonal Depression, *Canadian Journal of Psychiatry*, 47(4), 361-367.

Levitt, A. J., Boyle, M. H., Joffe, R. T., and Baumal, Z. (2000). Estimated Prevalence of the Seasonal Subtype of Major Depression in a Canadian Community Sample, *Canadian Journal of Psychiatry*, 45(7), 650-654.

Luo, F., Florence, C. S., Quispe-Agnoli, M., Ouyang, L., and Crosby, A. E. (2011). Impact of Business Cycles on US Suicide Rates, 1928-2007, *American Journal of Public Health*, 101(6), 1139-1146.

McMullan, M. (2006). Patients Using the Internet to Obtain Health Information: How This Affects the Patient–health Professional Relationship, *Patient Education and Counseling*, 63(1), 24-28.

Mersch, P. P. A., Middendorp, H. M., Bouhuys, A. L., Beersma, D. G., and van den Hoofdakker, R. H. (1999). Seasonal Affective Disorder and Latitude: A Review of the Literature, *Journal of Affective Disorders*, 53(1), 35-48.

Michalak, E. E., Wilkinson, C., Dowrick, C., and Wilkinson, G. (2001). Seasonal Affective Disorder: Prevalence, Detection and Current Treatment in North Wales, *The British Journal of Psychiatry*, 179(1), 31-34.

Ogburn, W. F., and Thomas, D. S. (1922). The Influence of the Business Cycle on Certain Social Conditions, *Journal of the American Statistical Association*, 18(139), 324-340.

Patel, V., and Saxena S. (2014). Transforming Lives, Enhancing Communities—Innovations in Global Mental Health, *New England Journal of Medicine,* 370(6), 498-501.

Parker, G., and Walter, S. (1982). Seasonal Variation in Depressive Disorders and Suicidal Deaths in New South Wales, *The British Journal of Psychiatry*, 140(6), 626-632.

Preis, T., Reith, D., and Stanley, H. E. (2010). Complex Dynamics of Our Economic Life on Different Scales: Insights from Search Engine Query Data, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 368(1933), 5707-5719.

Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M. R., et al. (2007). No Health Without Mental Health, *Lancet*, 370(9590), 859-877.

Ruhm, C. J. (1995). Economic Conditions and Alcohol Problems, *Journal of Health Economics*, 14(5), 583-603.

Sargent-Cox, K., Butterworth, P., and Anstey, K. J. (2011). The Global Financial Crisis and Psychological Health in a Sample of Australian Older Adults: A Longitudinal Study, *Social Science and Medicine*, 73(7), 1105-1112.

Suziedelyte, A. (2012). How Does Searching for Health Information on the Internet Affect Individuals' Demand for Health Care Services?, *Social Science and Medicine*, 75(10), 1828-1835.

Tefft, N. (2011). Insights on Unemployment, Unemployment Insurance, and Mental Health, *Journal of Health Economics*, 30(2), 258-264.

Thomas, D. S. (2015). *Social Aspects of the Business Cycle (RLE: Business Cycles),* Routledge.

Saarijärvi, S., Lauerma, H., Helenius, H., and Saarilehto, S. (1999). Seasonal Affective Disorders Among Rural Finns and Lapps, *Acta Psychiatrica Scandinavica*, 99(2), 95-101.

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., et al. (2013). Global Burden of Disease Attributable to Mental and Substance Use Disorders: Findings from the Global Burden of Disease Study 2010, *Lancet*, 382(9904), 1575-1586.

Yang, B. (1992). The Economy and Suicide, *American Journal of Economics and Sociology*, 51(1), 87-99.

Zivin, K., Paczkowski, M., and Galea, S. (2011). Economic Downturns and Population Mental Health: Research Findings, Gaps, Challenges and Priorities, *Psychological Medicine*, 41(7), 1343-1348.