An Infodemiology Approach to Assess the Impact of the Great Recession on Anxiety and Depression in France.

Ariadna Ramos-Gomez

Sciences Po

Aldo Perez-Escatel

Universidad Autonoma de Zacatecas

Elio Atenogenes Villaseñor-Garcia

Centro de Investigacion e Innovacion en Tecnologias de la Informacion y Comunicacion

Cesar Ramos-Remus (✉ r_amos@prodigy.net.mx)

Unidad de Investigacion en Enfermedades Cronico-Degenerativas, S.C. <https://orcid.org/0000-0002-4898-4219>

Research article

Keywords: France; Mental Health; Unemployment; Macroeconomic conditions; Health; Welfare.

Posted Date: November 26th, 2019

DOI: <https://doi.org/10.21203/rs.2.17816/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The global financial crisis of 2008 and subsequent sovereign debt crisis in Europe had a profound multidimensional impact on developed economies, including on the construct of health. Mental health is a key factor for productivity and an indicator of the general well-being of the population. France poses mental health and public policies characteristics that differentiated it from other countries of the European Union. This study aimed to assess whether depression and anxiety increased throughout the course of the 2008 debt crisis in France. **Methods:** This is an ecological study using an infodemiology approach. The search terms for depression and anxiety were extended with related words according to culture to form a distress index. Google Trends was used to mine the distress index search terms in a period before the debt crisis (2004) until 2018, using open source libraries, and pandas to merge information. Diverse economic- and employment-related datasets for France were used as covariates in multiple linear regression analyses, and the expected seasonal variability as a dummy variable. **Results:** The results showed that for each additional demand for unemployment benefits, search traffic for the distress index was expected to increase by 0.09% ($p < 0.001$, $r^2 0.89$), even after controlling for the expected seasonal variations in the distress index. However, there was no significant association with unemployment rate. **Conclusions:** Contrary to reports from other countries, the mild effect of demand for unemployment benefits on the distress index may be explained by aspects of social protection policies in France.

Background

There is growing evidence of a relationship between economic cycles and health. The global financial crisis of 2008 and the subsequent sovereign debt crisis in Europe had a profound impact on developed economies. Europe remains on the path to recovery a decade after the crisis, with policymakers and economists still figuring out ways to deal with its aftermath. However, evidence remains fragmented despite the growing interest in studying the effects of the recession on different health-related outcomes in Europe. For example, a recent systematic literature review [1] found that of 43 relevant studies published from January 2008 to December 2015 that analyzed the impact of the financial crisis on health in Europe, 73% had high risk for bias, 22% had a moderate risk for bias, and 5% (two studies) had a low risk for bias. This limited the conclusions that could be drawn from that review, meaning the results must be interpreted with caution.

Research on the health effects of previous recessions has presented inconsistent findings that differ based on the extent of the effects experienced by age group, sex, diseases, and indicators used to measure economic change. Some authors have argued that unemployment has positive effects on health [2,3,4,5]. Ruhm (2000) [2] found that economic crises may have been beneficial for health in the United States. He argued that recessions decreased mortality rates measured in terms of 10 specific illnesses, suicide being an important exception. Toffoluti and Suhrcke [5] used data for 23 European countries and found that overall, an increase in unemployment rates during the Great Recession had beneficial health effects across European countries, as measured by decreases in all-cause mortality, cardiovascular

diseases, cirrhosis and chronic liver disease-related mortality, motor vehicle accident-related mortality, and parasitic infection-related mortality. The suicide rate was again the sole exception, which increased by around 34% with every one percentage point increase in the standardized unemployment rate.

In contrast, other authors argued that economic downturns negatively affect health and well-being. Theodossiou et al. [6] investigated 13 European countries and found a strong, positive relationship between adverse economic conditions and the mortality rate for the same 10 illnesses studied by Ruhm [2]. Clark and Oswald [7] found that in the UK, unemployment was positively related to low disutility scores from being unemployed. This observed effect was stronger for those that had just lost their jobs (short-term unemployed) than those that had been unemployed for a longer period of time (long-term unemployed). Di Tella et al. [8] examined how survey respondents' reports of their well-being varied with differing levels of inflation and unemployment. They showed that within their sample of 12 European countries and the United States, people appeared to be happier when inflation and unemployment were low. Burgard et al. [9] showed that in the United States, involuntary job loss was associated with more depressive symptoms and significantly poorer overall self-rated health, even after controlling for numerous social background characteristics and baseline health.

Overall, the stress felt by individuals due to an economic crisis has been demonstrated to be strongly linked with their physical and mental health [2,6,10]. Mental health is a key factor for productivity and an important indicator of the overall well-being in a population. However, information on the relationship between economic cycles and specific domains of mental health is scarce in European countries. Yet, Internet search traffic may provide a unique insight into how people react to economic changes in different domains of their lives, including those related to their health [11,12, 13, 14]. Few studies have incorporated Google searches for such analyses. Choi and Varian [11] showed how data on Google searches can improve forecasts on economic time series and that simple seasonal models and fixed-effects models that included Google Trends variables tended to outperform models that excluded Google Trends by 5%–20%. Askatas and Zimmerman [12] found a strong correlation between keyword searches of depression and anxiety symptoms and unemployment rates in the G8 countries using monthly data. Tefft [15] incorporated data web searches from Google Insights for Search, and estimated the correlation between the Great Recession and indices for depression and anxiety in the United States. He found a positive relationship between the unemployment rate and the depression search index, and a negative relationship between initial unemployment insurance claims and depression and anxiety search indices. Evidence in the present paper is presented as being suggestive that these searches were meaningful representations of the intention to understand or seek treatment for symptoms of psychological distress that a person might be experiencing at the time of their search. Even without consensus on the relationship between recessions and physical health, it is generally agreed that economic downturns negatively affect citizens' mental health. In the case of France, Malard et al. [16] assessed the changes in mental disorders among the French working population using nationally representative and prospective data with diagnostic interviews for major depressive episodes (MDE) and generalized anxiety disorder (GAD). They found no changes in the prevalence of MDE and GAD in either sex in the working population during 2006–2010; the exception being an increase of 7.4% of these disorders among women working in

the public sector. The authors concluded that high levels of social protection in France may explain their results. Furthermore, Blasco and Brodaty [17] reported that unemployment had a negative effect on mental health among men but not among women in France using an instrumental variable and data regarding unemployment duration and the prevalence of depression and anxiety. They also found that psychological support for unemployed people can prevent mental health problems. Baker et al. [13] found that approximately 40% of survey respondents with Internet access reported using search engines for health information in 2001, including when searching for information on chronic conditions such as MDE. Lam-Po-Tang and McKay [14] showed that approximately 80% of patients attending two psychiatric institutes used the Internet to search for symptoms of mental disorders. Furthermore, there is a relatively new method of research on Internet search data called “infodemiology.” Infodemiology has been defined as “the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy” [18].

This study aimed to use an infodemiology approach to assess two dominions of mental health (depression and anxiety) during the 2008 debt crisis in France. Our premise was that the search traffic for depression- and anxiety-related terms increased during the worst years of the recession (2009–2014), implying a growing interest in the subject and therefore a possible decrease in the general mental health of the population.

Methods

- Study design and outcome variables:

This is an ecological study using an infodemiology approach. Following Tefft [15], we defined the terms *depression* and *anxiety* as the main outcomes of interest because they are commonly occurring non-psychotic mental disorders. However, by studying only those two specific words, Tefft excluded people that searched health information by entering symptoms or cultural words that would account for depression and anxiety and may differ by the subject’s education level and linguistic background. Examples of French terms for depression- and anxiety-related words include: *idées suicidaires* (suicidal thoughts), *désespoir* (despair), *fatigue* (fatigue) *avoir le cafard* (to feel down), and *coup de blues* (having the blues). For this reason, we created a “distress index,” which accounted for all the depression- and anxiety-related words that were searched in France during the period of interest.

- Distress index:

To calculate the distress index, we first defined depression- and anxiety-related vocabulary to be later used as query words in Google Trends using fastText, which is a library of word embeddings and text classification. Thirteen words were input with fastText classification: *dépression* (depression), *anxiété* (anxiety), *fatigue* (fatigue), *tristesse* (sadness), *isolement* (isolation), *mal-être* (ill-being), *insomnie* (insomnia), *désespoir* (despair), *morosité* (gloominess), *déprime* (depressed), *spleen* (spleen), *angoisse*

(anguish), and *tremblements* (trembling). These words were chosen because they were either symptoms of depression and anxiety or synonyms of the same. To estimate similar related concepts, we used the method of close neighbors, with a distance that measured the angle between the vectors calculated with fastText and those of the similar concepts. Each similar concept had a similarity index from 0 to 1 that was associated with one of the words mentioned above. The similarity index parameter was set to be greater than 0.5, meaning fastText would only produce similar concepts that had a similarity index greater than 0.5. The model produced a total of 317 words with a similarity index greater than 0.5. Of those 317 words, we excluded 206 words from the analysis because they had no relation to depression or anxiety. Finally, we chose 111 words from this analysis, and included several concepts derived from French informal language or slang: *idées suicidaires*, *idée suicidaire*, *avoir le cafard*, and *coup de blues*. In total, the distress index comprised 115 words related to depression and anxiety. These 115 words were then computed in Google Trends, and the results exported from the Google Trends health category by month and country level. Google Trends results were retrieved semi-automatically with the aid of API Pytrends written in Python. The API allowed us to take each word of the vocabulary and define a search query considering France our scope. By the use of Pandas library, we put together all the 115 retrieved records from Google trends. Finally, we calculated the average of the 115 records and grouped them by quarter. The range of the Google Trends results covered the first quarter of 2004 to the third quarter of 2018 (January 1, 2004, to October 30, 2018). We consider these average as a distress index for each quarter. These results formed the distress index, meaning the distress index was the average results of the 115 depression- and anxiety-related words calculated by quarter. In this analysis, we adopted a standard transformation of the exported search indices, such that the distress index across time had a mean of zero and a variance of one. This normalization allowed for an equal interpretation of coefficients across models. Figure 1 shows the evolution of the normalized distress index, with each vertical line representing the first quarter of each year (winter months). Figure 1 shows an increasing trend of depression searches over the studied period, which was particularly steep after 2008. This suggested increasing interest in depression- and anxiety-related terms, and may reflect psychological distress suffered at the time of the search that can be seen to have worsened after the beginning of the crisis. Figure 1 shows the same trend with a vertical line in the first quarter of each month. The evolution of the distress index shows that the search activity peaked in winter months (first quarter), and then decreased over the next three quarters. In addition, lower depression search activity was evident during the summer months. These seasonal cyclical patterns of the distress index over time are consistent with findings in the literature pertaining to seasonal affective disorder (SAD).

- Study design and outcome variables:

This is an ecological study using an infodemiology approach. Following Tefft [15], we defined the terms *depression* and *anxiety* as the main outcomes of interest because they are commonly occurring non-psychotic mental disorders. However, by studying only those two specific words, Tefft excluded people that searched health information by entering symptoms or cultural words that would account for

depression and anxiety and may differ by the subject's education level and linguistic background. Examples of French terms for depression- and anxiety-related words include: *idées suicidaires* (suicidal thoughts), *désespoir* (despair), *fatigue* (fatigue) *avoir le cafard* (to feel down), and *coup de blues* (having the blues). For this reason, we created a "distress index," which accounted for all the depression- and anxiety-related words that were searched in France during the period of interest.

- Distress index:

To calculate the distress index, we first defined depression- and anxiety-related vocabulary to be later used as query words in Google Trends using fastText, which is a library of word embeddings and text classification. Thirteen words were input with fastText classification: *dépression* (depression), *anxiété* (anxiety), *fatigue* (fatigue), *tristesse* (sadness), *isolement* (isolation), *mal-être* (ill-being), *insomnie* (insomnia), *désespoir* (despair), *morosité* (gloominess), *déprime* (depressed), *spleen* (spleen), *angoisse* (anguish), and *tremblements* (trembling). These words were chosen because they were either symptoms of depression and anxiety or synonyms of the same. To estimate similar related concepts, we used the method of close neighbors, with a distance that measured the angle between the vectors calculated with fastText and those of the similar concepts. Each similar concept had a similarity index from 0 to 1 that was associated with one of the words mentioned above. The similarity index parameter was set to be greater than 0.5, meaning fastText would only produce similar concepts that had a similarity index greater than 0.5. The model produced a total of 317 words with a similarity index greater than 0.5. Of those 317 words, we excluded 206 words from the analysis because they had no relation to depression or anxiety. Finally, we chose 111 words from this analysis, and included several concepts derived from French informal language or slang: *idées suicidaires*, *idée suicidaire*, *avoir le cafard*, and *coup de blues*. In total, the distress index comprised 115 words related to depression and anxiety. These 115 words were then computed in Google Trends, and the results exported from the Google Trends health category by month and country level. Google Trends results were retrieved semi-automatically with the aid of API Pytrends written in Python. The API allowed us to take each word of the vocabulary and define a search query considering France our scope. By the use of Pandas library, we put together all the 115 retrieved records from Google trends. Finally, we calculated the average of the 115 records and grouped them by quarter. The range of the Google Trends results covered the first quarter of 2004 to the third quarter of 2018 (January 1, 2004, to October 30, 2018). We consider these average as a distress index for each quarter. These results formed the distress index, meaning the distress index was the average results of the 115 depression- and anxiety-related words calculated by quarter. In this analysis, we adopted a standard transformation of the exported search indices, such that the distress index across time had a mean of zero and a variance of one. This normalization allowed for an equal interpretation of coefficients across models. Figure 1 shows the evolution of the normalized distress index, with each vertical line representing the first quarter of each year (winter months). Figure 1 shows an increasing trend of depression searches over the studied period, which was particularly steep after 2008. This suggested increasing interest in depression- and anxiety-related terms, and may reflect psychological distress

suffered at the time of the search that can be seen to have worsened after the beginning of the crisis. Figure 1 shows the same trend with a vertical line in the first quarter of each month. The evolution of the distress index shows that the search activity peaked in winter months (first quarter), and then decreased over the next three quarters. In addition, lower depression search activity was evident during the summer months. These seasonal cyclical patterns of the distress index over time are consistent with findings in the literature pertaining to seasonal affective disorder (SAD).

- Economic- and Employment-related datasets:

Summary statistics for unemployment data are shown in Table 1. We compiled a dataset of quarterly data on the unemployment rate and economic fluctuation variables. The sample time frame was from the first quarter of 2004 to the third quarter of 2018, accounting for the beginning, development, and spread of the sovereign debt crisis. Estimates of the seasonally adjusted unemployment rate, gross domestic product (GDP) growth, labor force and activity rate were taken from the OECD database by country level and quarter. The unemployment rate is the unemployment rate for all persons in France aged 15 to 74 years old and it is seasonally adjusted in order to remove influences of predictable seasonal patterns from month to month. GDP growth measures the change from the total value of final goods and services produced by a country during a quarter minus the value of imports; this indicator is seasonally adjusted and measures the percentage change from the same quarter of the previous year. Labor force, or currently active population, includes all persons who fulfill the requirements for inclusion among the employed (civilians and armed forces) or the unemployed, and is seasonally adjusted and measured in thousands. The activity rate is the ratio between the number of active persons (occupied labor force and the unemployed) and the corresponding total population. These data are internationally comparable by following the System of National Accounts. Unemployment benefits were broadly defined as the transfers and employment services made by state governments to individuals that were able to work, had lost their job, and were unable to find gainful employment. Unemployment benefits data were exported from the Pôle emploi (employment center) webpage. Pôle emploi is a French governmental agency established in 2009 that registers unemployed people and provides financial aid. The agency maintains an extensive dataset on the number of unemployment benefits demanded by age and region. These data were categorized by age groups, or demands from those: younger than 25 years; aged 25–50 years; over 50 years; and the overall demand for unemployment benefits in France.

Table 1. Summary Statistics for Unemployment Data

	Mean	Std.Dev.	Min	Max
Distress index	26.549	4.5	19.084	34.757
Normalized distress index	0	1	-1.659	1.824
Unemployment rate	9.012	.944	6.845	10.484
Unemployment benefits >50	861.753	327.587	482.9	1428.1
Unemployment benefits 25-50	2793.222	460.669	2054.9	3483.1
Unemployment benefits <25	693.059	81.897	520	810.2
Unemployment benefits all	4348.032	855.622	3059.7	5649.6
% change in purchasing power	.301	.504	-1.797	1.361
Labor force in thousands	28451.67	881.991	27012.05	29830.93
GDP Growth	1.255	1.467	-3.78	3.041
Activity rate	70.528	.751	69.487	72.026
Seasonal dummy variable	.254	.439	0	1
% Change in disposable income	.589	.589	-1.611	1.963

Notes: number of observations = 59. Each observation is at the country level by quarter, and the sample includes the years 2004–2018. Distress Index is the search scores from Google Trends, then normalized across the sample period. Unemployment data are drawn from the Organization for Economic Co-operation and Development (OECD). Data on percentage change in disposable income and percentage change in purchasing power were exported from the French National Institute of Statistics and Economic Studies (INSEE) database. Data on demands for unemployment benefits were exported from the French *Pôle Emploi* dataset.

Percentage change in disposable income and percentage change in purchasing power were exported from the French National Institute of Statistics and Economic Studies (INSEE) databases. Percentage change in disposable income is the change from quarter to quarter on the income remaining after deduction of taxes and social security charges. The percentage change in purchasing power is the value of a currency expressed in terms of the amount of goods and services one unit of money can buy. It is important because it considers inflation, and the decrease of goods and services an agent can buy as it increases.

- Exploratory analysis:

We created a correlation matrix as the first step to prove a possible correlation between the variables, support the need for further analysis, and choose which variables seemed most appropriate for inclusion in the empirical strategy. As shown in Table 2, our variable of interest (normalized distress index) was positively correlated with the unemployment rate and demand for benefits by age. These results indicated that if the search traffic for the distress index grew, the unemployment variables would also increase. As expected, it was also observed that the labor force was positively correlated with the unemployment variables (unemployment benefits and unemployment rate). The GDP growth variable did not appear to have any strong correlations with any of the study variables. Finally, as shown in Table 2, there was a problem of collinearity between the variables and unemployment benefits by age. Therefore, we used the overall demand for benefits and not the variable categorized by age in the statistical analyses. The correlation matrix suggested that there was a relationship between our variables, which supports a more thorough study of the same.

Table 2. Correlation Matrix for study variables

	CDI	AR	DI	GDP	LF	PP	UB	UB25	UB25-50	UB50	UN
CDI	1.00										
AR	-0.37	1.00									
DI	-0.38	0.90	1.00								
GDP	0.49	-0.13	-0.09	1.00							
LF	-0.32	0.95	0.93	-0.11	1.00						
PP	0.76	-0.19	-0.26	0.07	-0.14	1.00					
UB	-0.25	0.94	0.89	0.07	0.90	-0.13	1.00				
UB25	-0.38	0.80	0.77	-0.11	0.73	-0.17	0.88	1.00			
UB25-50	-0.23	0.92	0.87	0.10	0.88	-0.11	1.00	0.89	1.00		
UB50	-0.24	0.95	0.91	0.07	0.94	-0.13	0.99	0.81	0.98	1.00	
UN	-0.31	0.75	0.73	-0.09	0.72	-0.09	0.85	0.95	0.85	0.78	1.00

CDI: percentage change in disposable income; AR: activity rate, DI: normalized distress index, GDP: GDP growth; LF: labor force in thousands; PP: percentage change in purchasing power; UB: unemployment benefits for all; UB25: demand for unemployment benefits from those younger than 25; UB25-50: demand for unemployment benefits from those aged 25-50; UB50: demand for unemployment benefits from those older than 50; UN: unemployment rate.

We adopted two models to estimate the effects of unemployment variables on the distress index, building on the work by Tefft [15]. In the basic model 1, we performed linear regression using the following empirical model:

$$\text{DistressIndex}_{tq} = \alpha + \beta \text{UnemploymentBenefitsAll}_{tq} + \delta \text{GDPGrowth}_{tq} + \gamma \text{PP}_{tq} + \gamma_q + \varepsilon_{tq} \quad (1)$$

where $\text{DistressIndex}_{tq}$ is the normalized “distress index” from Google Trends in France during the year t and quarter q ; $\text{UnemploymentBenefitsAll}_{tq}$ is the overall demand for unemployment benefits in France, during the year t and quarter q ; GDPGrowth_{tq} is the rate of GDP growth in each quarter with respect to the last year’s quarter; PP_{tq} is the percentage change in purchasing power (PP) for year t and quarter q ; and γ is a seasonal dummy variable that takes the value 1 for each first quarter (winter months), and the value 0 for the other three quarters. The dummy variable was important given the seasonal variations in mood and psychological distress due to seasonal effects, a phenomenon that has been highly documented [19, 20, 21]. Our coefficient of interest was β , which identifies the effects of unemployment fluctuations on mental health outcomes. We chose to only use the variable unemployment benefits and not the unemployment rate because, as seen in the matrix, both variables were highly correlated. Therefore, including both variables might have caused a serious problem of collinearity. The same reasoning was applied to the other variables that were not included in the model.

In model 2, we added an autoregressive model of lag 1 (AR(1)) to correct the model from a serial correlation problem in the residuals. This model predicted future behavior based on past behavior. The simple regression model with AR(1) structure is:

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t; \mu(\varepsilon_t | \varepsilon_1, \dots, \varepsilon_{t-1}) = \alpha \varepsilon_{t-1}$$

Using a filtering transformation, $V_t = Y_t - \alpha Y_{t-1}$ and $U_t = X_t - \alpha X_{t-1}$, we regressed V_t on U_t to get AR(1) regression estimates. Adding an autoregressive model of lag 1 (AR(1)) gave the following and final model:

$$\text{DistressIndex}_{tq} = \alpha + \beta \text{UnemploymentBenefitsAll}_{tq} + \gamma_q + \varepsilon_{tq} \quad (2)$$

where $\text{DistressIndex}_{tq}$ is the normalized “distress index” from Google Trends in France during the year t and quarter q ; $\text{UnemploymentBenefitsAll}_{tq}$ is the overall demand for unemployment benefits in France during the year t and quarter q ; an γ is a seasonal dummy variable with the value 1 for the winter months (first quarter of each year) and the value 0 for the other three quarters. We excluded the variables GDP growth and the percentage change in PP because after correcting the issue of serial correlation, these variables were not significant. We then regressed the subtraction of the normalized distress index from the previous period ($Y_t - \alpha Y_{t-1}$) on the subtraction of the variable of unemployment benefits from the previous period ($U_t = X_t - \alpha X_{t-1}$) to get the AR(1) estimates. Accounting for a possible serial correlation meant that our final model (2) was significant and unbiased.

Results

Table 3 shows the correlation between the demand for unemployment benefits and the distress index. The results were similar to previous studies that showed a positive correlation between unemployment and psychological distress. In particular, for each additional demand for unemployment benefits, the search traffic of the index was expected to increase by 0.10%, an effect that was significant at the 1% level. Furthermore, GDP growth and the percentage change in PP had a negative correlation with the distress index, an effect that was statistically significant at the 5% level. This also appeared to be predictable given the fact that as GDP and PP increased in a certain population, the stress level decreased, which may have caused a decline in depression- and anxiety-related searches. As shown by Easterlin [22], income levels are directly related to individual happiness and satisfaction to a certain extent. The dummy variable accounting for the seasonal effect of psychological distress-related searches had a positive correlation with our dependent variable, an effect that was statistically significant at the 5% level and had a much higher coefficient than the unemployment variables. Nevertheless, the model appeared to have a problem of serial correlation in the residuals, as shown by the Durbin-Watson statistic. Normal Durbin-Watson values vary from 1.5 to 2.5; values outside this range show a possible serial correlation. In particular, those below 1 and above 3 suggest a definitive problem of serial correlation. The model had a Durbin Watson value below 1.

Table 3. Main Results: First Model (1).

Variable	Coefficient	Std. error	t-statistic	Prob.
C	−4.371144***	0.285625	−15.30377	0.0000
UB	0.001033***	6.27E-05	16.47597	0.0000
GDPGROWTH	−0.095742**	0.036367	−2.632676	0.0110
PP	−0.264234**	0.106406	−2.483255	0.0162
D1	0.310006**	0.120841	2.565400	0.0131
R-squared	0.848112	Mean dependent variable −1.64E-08		
Adjusted R-squared	0.836861	SD dependent variable 1.000000		
SE of regression	0.403905	Akaike info criterion 1.105665		
Sum squared residual	8.809528	Schwarz criterion 1.281728		
Log likelihood	−27.61712	Hannan-Quinn criterion 1.174393		
F-statistic	75.38103	Durbin-Watson statistic 0.958748		
Prob (F-statistic)	0.000000			

SD, standard deviation; SE, standard error; GDP, gross domestic product; PP, purchasing power; C, constant; UB, unemployment benefits; D1, dummy variable accounting for winter months.

Notes: Included observations = 58 after adjustments. Method: Least Squares. Convergence achieved after 6 iterations.

** Significant at 5%.

*** Significant at 1%.

Table 4 presents the results of model 1, after correcting the serial correlation problem with an autoregressive model of lag 1. The Durbin Watson statistic fell within the normal parameter. However, the GDP and PP variables were not significant

Table 4. Main results: Model 1, after correcting for serial correlation

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.176204***	0.530748	-7.868532	0.0000
UB	0.000974***	0.000119	8.176597	0.0000
GDPGROWTH	-0.045510	0.057446	-0.792229	0.4318
PP	-0.112753	0.084826	-1.329227	0.1896
D1	0.262098***	0.076500	3.426110	0.0012
AR(1)	0.574145***	0.118121	4.860648	0.0000
R-squared	0.900444	Mean dependent var	0.024649	
Adjusted R-squared	0.890871	S.D. dependent var	0.990489	
S.E. of regression	0.327205	Akaike info criterion	0.701235	
Sum squared resid	5.567265	Schwarz criterion	0.914384	
Log likelihood	-14.33580	Hannan-Quinn criter.	0.784260	
F-statistic	94.06373	Durbin-Watson stat	2.390839	
Prob(F-statistic)	0.000000			
Inverted AR Roots	.57			

SD, standard deviation; SE, standard error; GDP, gross domestic product; PP, purchasing power; C, constant; UB, unemployment benefits; D1, dummy variable accounting for winter months.

Notes: Included observations = 58 after adjustments. Method: Least Squares. Convergence achieved after 6 iterations.

** Significant at 5%.

*** Significant at 1%.

Table 5 shows the final results with model 2, after correcting for the serial correlation problem and removing the variables that were not significant. Inverted AR roots were within normal parameters and the R-squared value increased to 89%, meaning almost 90% of the variance in the dependent variable could be explained by the model. This final model estimated that for each additional demand for unemployment benefits, the search traffic for the distress index was expected to increase by 0.09%, which was significant at the 1% level. The seasonal dummy variable showed an increase in psychological distress-related searches in winter, specifically, an increase of 26.3% during the first quarter of the year, which was significant at the 1% level.

Table 5. Main results: Model 2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.186723***	0.620581	-6.746456	0.0000
UB	0.000957***	0.000137	6.969841	0.0000
D1	0.262938***	0.074250	3.541224	0.0008
AR(1)	0.632485***	0.101357	6.240174	0.0000
R-squared	0.896130	Mean dependent var	0.024649	
Adjusted R-squared	0.890360	S.D. dependent var	0.990489	
S.E. of regression	0.327970	Akaike info criterion	0.674685	
Sum squared resid	5.808488	Schwarz criterion	0.816785	
Log likelihood	-15.56588	Hannan-Quinn criter.	0.730036	
F-statistic	155.2940	Durbin-Watson stat	2.476493	
Prob(F-statistic)	0.000000			
Inverted AR Roots	.63			

SD, standard deviation; SE, standard error; C, constant; UB, unemployment benefits; D1, dummy variable accounting for winter months; AR(1), autoregressive model of lag 1.

Notes: Included observations = 58 after adjustments. Method: Least Squares. Convergence achieved after 5 iterations.

** Significant at 5%

*** Significant at 1%

Finally, we ran model 2 with the unemployment rate (excluding the unemployment benefits demand) as the independent variable. Surprisingly, the model was not significant. This suggested that the unemployment rate and the demand for unemployment benefits were different, with the latter capturing a loss in income. The unemployment rate may capture people who do not have a job, but does not necessarily capture people in need of income. The people who demand unemployment benefits are, by definition, in need of income; therefore, this variable remained significant because it assessed the relationship between income necessities and psychological distress-related searches.

Discussion

To our knowledge, this study was the first ecological study to use an infodemiology approach to explore the changes in depression and anxiety during the sovereign debt crisis in France, as captured by Google Trend searches. The objective of this study was to offer a comprehensive approach to assessing the impact of the sovereign debt crisis on mental health in France. We used extensive Google search data that accounted for the words *depression* and *anxiety* as well as for the terms associated with them. We confirmed previous findings where the three models yielded positive unemployment coefficients, meaning that the demand for unemployment benefits was positively and significantly associated with the distress index. At first glance, this significant association may be explained in terms of the crisis causing more people to be interested in symptoms related to depression and anxiety because more people were experiencing psychological distress of some kind. However, at the same time, the strength of the effect of the unemployment variable on mental health was minimal when compared with the effect found by SAD. Therefore, even if unemployment may be damaging to mental health, its effect appeared to have had a minor impact in France. This finding was consistent with broad research during previous recessions that explored the positive effects of social protection, where social protection expenditures appeared to help countries counteract deteriorating mental health (e.g., via unemployment insurance). For example, O'Campo et al. [23] showed that psychological distress among the unemployed and employed was impacted by unemployment insurance schemes and could moderate harmful consequences of unemployment. Unemployment benefits tend to smooth patterns of consumption during economic downturns.

The recent global crisis makes unemployment insurance an increasingly crucial support program for the growing number of unemployed people. In the French universal health system, citizens have access to healthcare and psychological help. A World Health Organization report in 2000 [24] recognized France's healthcare system (including mental health assessment) as number one out of 191 countries. The free universal access to psychologists, psychiatrists, and medicines as well as unemployment insurance

schemes in France may explain our results, and highlight the importance of the welfare state at the aggregate level.

Despite the robust results presented, certain limitations of this study need to be considered. First, the results of this study do not exclude the likelihood that there may be adverse long-term mental health effects or the possibility that aggregate effects differ from the same effects at the individual level. It might be that certain parts of the population were worse off in terms of health losses, while others were better off. For example, if an increase in unemployment within a country (in this case France) is found to be correlated with an increase in mental health-related problems or even suicide risk, we cannot exclude the possibility that within that country, those people that were worse off in terms of mental health were actually the ones who were employed (not unemployed), but their living conditions had stalled.

Second, this study was based on the assumption that Google Trends searches provide a good proxy of the state of health within a certain group, an assumption that has been treated largely in academic literature [11,12,13,14]. However, Google Trends has its own limitations, and it might be that increasing Google searches on a certain health topic are not attributable to an increase in the prevalence of that disease, but rather an increase in reporting of the illness because of changes in the number of times individuals searched using the Internet. This implies that Google searches for depression- and anxiety-related terms increased over the studied period, not because the interest in depression and anxiety increased, but because Google's popularity increased and searching information on the Internet became more conventional. However, Google continues to improve its services, and therefore continually changes its data-generating processes and search algorithms [25,26]. In addition, Google Trends is not designed for health research, and the present results may not be reproducible if searches are performed using different dates. There are also other platforms with big data for social media analytics, such as Facebook Insights and Twitter Analytics. Although each platform has different profiles of users and contents and were created for business, they can also be used in infodemiology.

Third, the study subject was the searches themselves, not the people who performed them. Consequently, it is not possible to determine the number of individuals who performed the searches or their sociodemographic characteristics. In addition, such data does not directly control for enduring individual characteristics that may cause discrimination among the pool of employed individuals and those suffering from psychological distress. This may suggest a possible "reverse-causation" problem, where people suffering from depression and anxiety may lower their performance at work, thereby damaging their status within the workforce. It may be that unemployment (and especially loss in income) affects mental health, but it also may be possible that a mental health problem such as depression can cause some people to lose their job.

Finally, this type of study precludes statements on causality or associations; data interpretation is limited to assumptions. The correlation between unemployment variables and Google searches identified in this study is intended to be informative. More research is necessary to identify any causal relationships, especially between changes in the unemployment rate and mental health.

Conclusion

The Internet as a source of data for the social sciences, and for health-related subjects in particular, is a novel and efficient way to approach research studies. Google Trends represents an innovative tool of involuntary panel surveys and provides real-time reporting of reality. Despite the possible loss of valuable information caused by aggregation, it has been shown to provide researchers with a unique insight into large-scale processes in an easy and rapid manner.

We presented Google search data using an infodemiology approach to assess the impact of the sovereign debt crisis on mental health in France. By creating a distress index that accounts for depression and anxiety and their respective related terms, we found that unemployment variables had a positive, but minor impact on mental health in France. The strength of this study is the large Google searches dataset used to create the distress index, accounting for the words *depression* and *anxiety* and related terms. We suggest that the generous French unemployment insurance scheme and efficient universal healthcare system prevented worse psychological distress outcomes.

Given the high governmental deficit and debt levels of France, there have recently been concerns with respect to the social welfare safety nets, including the extent of healthcare and the amount of income transfers to the unemployed. Policymakers are therefore debating whether the French welfare state is too generous towards its people and whether it is still feasible to fund. The debate has excluded health-related outcomes and has focused on macroeconomic conditions. Policymakers should consider physical and mental health when enacting new reforms. Furthermore, the French welfare state, unemployment insurance, and universal healthcare, should remain untouched. However, more research is necessary to identify the causal relationship between unemployment variables, changes in macroeconomic conditions, and mental health. Mental health is a key factor for a productive population, and should be at the center of governmental agendas.

Abbreviations

MDE = Major depressive episode

GAD = generalized anxiety disorder

SAD = seasonal affective disorder

GDP = gross domestic product

OECD = Organization for Economic Co-operation and Development

INSEE = French National Institute of Statistics and Economic Studies

Declarations

- Ethics approval and consent to participate: Not applicable. This research does not involve human subjects, human material or human data. It analyses datasets of public domain. For this reason, the approval of some ethics committee is not required.
- Consent for publication: Not applicable
- Availability of data and materials: Supporting data used and analyzed in this manuscript is of public domain.
- Competing interests: The authors declare that they have no competing interests.
- Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
- Authors' Contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by ARG, APE, EVG and CRR. The first draft of the manuscript was written by ARG and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.
- Acknowledgements: We thank the members of the INGEOTEC research group (www.ingeotec.mx), particularly Daniela Moctezuma, Mario Graff, Eric Tellez and Sabino Miranda, for their recommendations and advice on the use of fastText. We thank Audrey Holmes, MA, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.
- Authors' information: We are grateful to note that this work is the condensed and edited version of an academic thesis presented by ARG in June 2019 at Sciences Po School of Public Affairs as part of the procedures to obtain her master's degree in economics.

References

- 1.- Parmar, D., Stavropoulou, C., Ioannidis J.P. 2016. Health outcomes during the 2008 financial crisis in Europe: systematic literature review. *British Medical Journal* Sept 6. 354,i4588. doi: 10.1136/bmj.i4588.
- 2.- Ruhm, C.J., 2000. Are recessions good for your health? *Quarterly Journal of Economics* 115 (2), 617–650. doi.org/10.1162/003355300554872.
- 3.- Ruhm, C.J. 2003. Good times make you sick. *Journal of Health Economics* 22(4), 637–658. doi: 10.1016/S0167-6296(03)00041-9.
- 4.- Ruhm, C.J., 2005. Healthy living in hard times. *Journal of Health Economics* 24(2), 341–363. DOI: 10.1016/j.jhealeco.2004.09.007.
- 5.- Toffolutti, V., Suhrcke, M. 2014. Assessing the short term health impact of the Great Recession in the European Union: A cross-country panel analysis. *Preventive Medicine* 64, 54-62. doi: 10.1016/j.ypmed.2014.03.028.
- 6.- Theodossiou I., Economou A., Nikolaou, A. 2008. Are recessions harmful to health after all?: Evidence from the European Union. *Journal of Economic Studies* 35, 368-384. doi:/10.1108/01443580810903536.

- 7.- Clark, A.E., Oswald, A.J. 1994. Unhappiness and Unemployment. *The Economic Journal* 104(424), 648-659. DOI: 10.2307/2234639
- 8.- Di Tella, R., MacCulloch, R. J., Oswald, A. 2001. Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *The American Economic Review* 91(1), 335-341. Available from <https://www.jstor.org/stable/2677914>
- 9.- Burgard, S. A., Brand, J. E., House, J. S. 2007. Toward a Better Estimation of the Effect of Job Loss on Health. *Journal of Health and Social Behavior* 48(4), 369–384. doi:10.1177/002214650704800403
- 10.- Whitehead, B. R., & Bergeman, C. S. (2015). The effect of the financial crisis on physical health: Perceived impact matters. *Journal of Health Psychology*, 22(7), 864–873.
- 11.- Choi, H., Varian, H. 2012. Predicting the Present with Google Trends. *Economic Record* 88(1), 2–9. doi:10.1111/j.1475-4932.2012.00809.x.
- 12.- Askitas, N., Zimmermann, K. F., 2011. Health and Well-Being in the Crisis. IZA Institute for the Study of Labor. Discussion Paper No. 5601. [Internet]. [accessed 2019 September 9]. Available from: <http://ftp.iza.org/dp5601.pdf>
- 13.- Baker, L., Wagner, T.H., Singer, S., Bundorf, M.K. 2003. Use of the Internet and Email for health care information: results from a national survey. *Journal of the American Medical Association* 289 (18), 2400–2406. DOI: 10.1001/jama.289.18.2400
- 14.- Lam-Po-Tang, J., McKay, D. 2010. Dr Google, MD: a survey of mental health-related Internet use in a private practice sample. *Australasian Psychiatry: bulletin of Royal Australian and New Zealand College of Psychiatrists* 18 (2), 130–133. doi: 10.3109/10398560903473645.
- 15.- Tefft, N. 2011. Insights on unemployment, unemployment insurance, and mental health. *Journal of Health Economics* 30 (3), 258–264. doi:10.1016/j.jhealeco.2011.01.006
- 16.- Malard L, Chastang JF, Niedhammer I. (2015). Changes in major depressive and generalized anxiety disorders in the national French working population between 2006 and 2010. *Journal of Affective Disorders* 178, 52-9. doi:10.1016/j.jad.2015.02.027.
- 17.- Blasco, S., Brodaty, T. (2016). Chômage et santé mentale en france. *Économie et statistique* 486-487, 17–44. [Internet]. [accessed 2019 September 9]. Available from: (https://www.persee.fr/doc/estat_0336-1454_2016_num_486_1_10688).
- 18.- Eysenbach G, (2006). Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annual Symposium Proceedings Archive* 2006, 244–248. [Internet] [cited 2019 September 19]. Available from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1839505/>

- 19.- Lam, R. W., Tam, E. M., Yatham, L. N., Shiah, I.-S., & Zis, A. P. (2001). Seasonal depression: the dual vulnerability hypothesis revisited. *Journal of Affective Disorders* 63(1-3), 123–132. doi:10.1016/s0165-0327(00)00196-8.
- 20.- Martiny, K., Lunde, M., Uden, M., Dam, H., & Bech, P. 2005. Adjunctive bright light in non-seasonal major depression: results from clinician-rated depression scales. *Acta Psychiatrica Scandinavica* 112(2), 117–125. doi:10.1111/j.1600-0447.2005.00574.x
- 21.- Pinchasov, B. B., Shurgaja, A. M., Grischin, O. V., Putilov, A. A. 2000. Mood and energy regulation in seasonal and non-seasonal depression before and after midday treatment with physical exercise or bright light. *Psychiatry Research*, 94(1), 29–42. doi:10.1016/s0165-1781(00)00138-4.
- 22.- Easterlin, R. A. (2001). Income and Happiness: Towards a Unified Theory. *The Economic Journal*, 111(473), 465–484. doi:10.1111/1468-0297.00646.
- 23.- O'Campo, P., Molnar, A., Ng, E., Renahy, E., Mitchell, C., Shankardass, K., St John A., Bamba C., Muntaner, C. 2015. Social welfare matters: A realist review of when, how, and why unemployment insurance impacts poverty and health. *Social Science & Medicine*, 132, 88–94. doi:10.1016/j.socscimed.2015.03.025
- 24.- WHO (2000). Rapport sur la santé dans le monde. Pour un système de santé plus performant. [Internet]. [accessed 2019 September 9] Available from: https://www.who.int/whr/2000/en/whr00_fr.pdf
- 25.- Lazer D, Kennedy R, King G, Vespignani A. 2014. Big data. The parable of Google flu: traps in big data analysis. *Science* 343,1203–1205. doi:10.1126/science1248506

Figures

Figure 1

Normalized Distress Index for France 2004-2018 representing winter months. Each first quarter of each year is represented with a line. For simplicity, years are written with their last two digits. The first quarter represents the winter months (January 1– March 31) of each year, from January 1, 2004 to September 30, 2018. The Normalized Distress Index has a mean of zero and a variance of one.