Analysis of Google search trends for measures of magical thinking in response to economic shocks

Raphael De Gottardi, Michael Wenger

July 14, 2024

1 Introduction

In this project, we aimed to analyze the behavior of the population when involved in economic shocks. For our evaluation, we set the focus on behavior related to 'magical thinking'. We hypothesized that people tend to have a higher interest in 'magical thinking' in times of uncertainty and economic weakness. We first wanted to check whether a correlation between interest in magical thinking and economic crises can be established, and then to see whether even a causality can be found. If this effect actually exists, it would be interesting to know, as it could be counteracted with preventive measures, for example, and people could be provided with alternative mental support during times of crisis.

We took a closer look at one specific economic shock: the Brexit referendum (June 23, 2016). For quantifying the interest in magical thinking, we use Google Trends (GT) as our source. Research using Google Trends has increased dramatically in the last decade. Proving that it is a powerful tool to analyze, describe, and diagnose research trends. It has already been used to track and predict flu outbreaks for example by correlating search queries with official flu data and has also been used to monitor the spread of COVID-19. Furthermore it has found an application in the field of economics where it was a useful tool to predict stock market movements based on search volumes for specific terms and to analyze trends in consumer behavior and changes in demand for products and services ([1], [2], [4]). However it comes with its limitations which will be discussed.

Our analysis is based on the assumption that interest is reflected in Google searches. We used the time series data of relative search volume (RSV) to quantify interest and compared it to indicators which represent the impact of the economic shock (i.e. the exchange rate of the GBP).
To substantiate magical thinking, we generated a list of 15 keywords related to the topic. For each of them, we downloaded the RSV time series for a given time period for analysis.

After having done our exploratory data analysis we started by investigating the Spearman Rank Correlation Coefficient of the RSV trajectories with the indicator values. Furthermore we used an approach based on the mean interest values before and after to look for trends in the data and compared those between more (UK) and less (US) affected regions. These two methods gave a good first insight in the data, given the constraints of GT data.

We continued by exploring more sophisticated methods. The diff-in-diff method which is a very common technique in econometrics based on linear regression of the data for more and less affected regions. Moreover we also explored the synthetic control method which uses multiple other data sources (in our case data from many different countries) to generate a better control variable. It is then compared to the obtained data to see if the difference is statistically relevant. Regarding the Brexit referendum we did not find any evidence for an increase in interest in response to the Brexit referendum based on the observed keywords.

2 Literature Review

Magical thinking as such is not well defined and only few sources have written on related topics. Sources we found focused more on the psychoanalytic aspects of magical thinking and are thus more related to the field of psychology, see [3] for example.
Although Google Trends has been developed mainly for journalism [8], using Google Trends as a source of data for scientific analysis has been done. Amaryllis Mavragani et al. [5] give an overview of methods, tools, and statistical approaches of papers using Google Trends data. They describe the increase in usage of such data and provide a summary of used methods. A paper with a similar goal to ours was published in 2021 by Jadwiga Hamulka et al [7]. It investigates interest in dietary supplements during the COVID-19 outbreak. As a measure of interest in such supplements, they use Google Trends data. More precisely they examined changes in interest in nutrients, bio-active compounds, and herbs in Poland in response to the covid pandemic. They found correlations between such search terms and their indicator variables.

3 Data and Summary Statistics

3.1 Google Trends data

Google saves the time history of various search queries and makes it available in real-time in the form of Google Trends. In Google Trends one can view the most popular search queries at the moment and one can also search for specific keywords. This data can be used as a measure that tells us about the interests of people who are living in a specific region during a specific time. Data is available since 2004 and as improvements have been made, more recent data is more accurate. It has already successfully been used for predict disease outbreaks ([1]), tourism flows ([6]) and trading behaviour in financial markets ([2]).

Since we want to analyze people's search behaviour we used Google Trends data as our data source. The data obtained from Google Trends can be downloaded for free in the form of a csv-file. When searching for a particular search term during a period, the resulting data is divided into time intervals which always contain a number between 0 and 100. This number is referred to as the relative search volume (RSV) which is always given in percent. It is normalized in a way that an amount of 100 corresponds to the maximum amount of queries corresponding to this search term which occurred during the selected period. An amount of 0 respectively means there were almost no searches. It is also possible to compare different search terms with each other. In this case, all values are subjected to a different normalization. Which means that each search term that is included in the search will have values between 0 and 100. Searches can also be grouped by location, time, and categories like sports, sciences or finances.

We will illustrate the possibilities of this data source with the use of a short example. A plot of the acquired data can be seen in Figure 1. In this example we display the RSV of the search terms "Superstition" and "Astrology" in the UK. We collected the data starting from June 23 in 2015 up to June 23 in 2017 which means we enclose the Brexit referendum by one year. There is an interval of one week between the respective RSV values.

It is evident that the data is quite noisy, this has to do with the rather low absolute interest in such topics, which leads to high variance in interest over short time periods. We provide some summary statistics for the first 5 terms of the data gathered in Table 1. They are representative for the dataset, more details can be found in the source data.

	Astrology	**Horoscope**	**Fortune telling**	**Psychic**	**Tarot reading**
count	105.0	105.0	105.0	105.0	105.0
mean	59.2	60.9	29.0	75.0	77.1
std	8.1	6.9	26.6	7.4	11.2
min	46.0	51.0	0.0	62.0	53.0
25%	54.0	57.0	0.0	70.0	68.0
50%	58.0	60.0	36.0	74.0	79.0
75%	62.0	63.0	47.0	79.0	85.0
max	100.0	100.0	100.0	100.0	100.0

Table 1: Statistics for the first five keywords

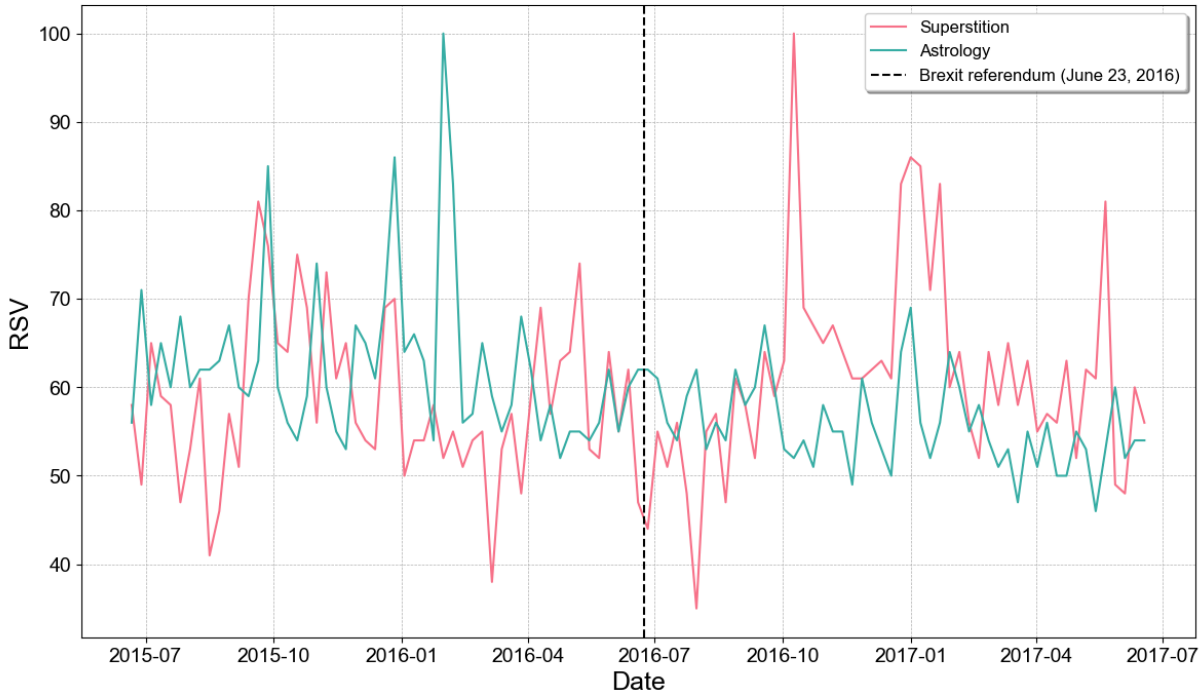


Figure 1: Trends over time for various keywords

We collected data for the time interval June 23 in 2015 to June 23 in 2017 for 6 different countries (two letter codes: UK,US,CA,AU,NZ,ZA) which correspond to the largest countries where English is a main national language. Only South Africa (ZA) had to few datapoints for 3 keywords, else we were able to obtain all the data needed. As a sidenote, downloading large amounts of data from Google Trends is nontrivial, the procedure we used is explained in the readme-file of the Github repository.

3.2 Sources for indicator data

To represent the impact of the economic shock, we used the exchange rate between USD and GBP as indicator. Historical data about exchange rates can be easily accessed and downloaded on `https://www.investing.com/` [10], where we used especially the closing price for one GBP in USD for each day. The time evolution of this quantity during our chosen time interval can be seen in Figure 2.

Figure 2: Closing price around Brexit referendum of GBP in USD

4 Methods

We started by generating a list of keywords related to magical thinking. From this list of search terms, we determined the RSV values from Google Trends. We then analyzed these values for correlation with the indicator values (the exchange rate of GBP and USD) and then, in another approach, applied the differences in differences method. We evaluated the correlation with the help of the Spearman Rank Correlation Coefficient. Lastly we used synthetic controls as a last method to analyse our data.

The data analysis has been done using Python jupyter notebook and libraries which will be mentioned below. The code which can be used to reproduce all our result can be found in our Github repository which includes an environment.yml file and also contains additional information in its readme-file.

4.1 Choice of search terms

In order to get a list with different search terms for which we could analyse the corresponding RSV values we used ChatGPT4.0 [9]. The prompt used for the keyword generation can be found in the appendix as well as in the Github repository.

The list of keywords which resulted is the following:

- Astrology
- Horoscope
- Fortune telling
- Psychic
- Tarot reading
- Numerology
- Palm reading
- Clairvoyant
- Crystal ball
- Magic spells
- Occult
- Spiritual healing
- Witchcraft
- Mediumship
- Superstition

4.2 Spearman's Rank Correlation Coefficient

In order to see whether a correlation between the RSV values of the search terms of our list and the price of GBP in USD is present or not we calculated the Spearman Rank Correlation Coefficient for each search term. The Spearman Rank Correlation Coefficient can be used to evaluate the strength and direction of the monotonic relationship between two variables which is exactly what we wanted to do. A monotonic relationship between two variables is present if either an increase of one value means also an increase of the other value or an increase of one value leads to a decrease of the other value. It is important to note here that this relationship does not have to be linear.

The definition of the Spearman rank correlation coefficient ρ_S can be seen in equation 1, where d_i is the difference between the ranks of the corresponding variables and n is the number of observations. This means that after ranking the observations the coefficients can be easily calculated. [11]

$$\rho_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{1}$$

Out of the definition follows that $\rho_S \in [-1, 1]$ where a value of 1 means that a perfect association of ranks is present and a value of -1 indicates a perfect negative association of ranks. The closer ρ_S is to zero, the weaker is the monotonic relationship between the two variables. If a positive or negative correlation exists, it must be examined for statistical significance. This can be done by choosing a null hypothesis and calculating the corresponding P-value. In our case the null hypothesis was that no monotonic relationship is present at all between our two variables. A very small P-value means then that under the null hypothesize it is very unlikely to observe such a high (or low) correlation coefficient which would then lead to rejection of the null hypothesize. [11]

4.3 Differences in differences

The differences-in-differences method (DiD) is a technique used in econometrics and other social sciences to estimate the causal effect of a treatment or intervention. It compares the changes in outcomes over time between a group that is exposed to the treatment (treatment) and a group that is not (control). In our case this would be a group affected by the Brexit referendum and the related economic shock and a group which is not. The key assumption is that, in the absence of the treatment, the difference in outcomes between the two groups would have remained constant over time. By observing the differences in outcomes before and after the treatment for both groups, the effect of the treatment could be isolated. However this method turns out to be tricky when working with Google Trends data because of the way it is normalised. This is best illustrated by the fact that the RSV is capped at 100, however the DiD method could expect values which are above 100.

4.3.1 Exploring means

Nevertheless, we provide the mean values before and after the Brexit referendum which give a good impression of how the Brexit referendum could have affected interest (in magical thinking). The mean value helps neglecting the large fluctuations that the data has which makes the analysis hard. Also we provide the same plot for data from the US which we use as control.

4.3.2 Diff-in-diff

Although the structure of Google Trends data is not well suited for the diff in diff method we can assume that this has only a minor effect on the outcome. This assumption allowed us to explore this method and get to know the procedure while being able to get results for our data which help interpreting it. We therefore quickly introduce the method:

The diff-in-diff approach assumes that in the absence of the treatment, the difference in outcomes between the treatment and control groups would have remained constant over time. By comparing the changes in outcomes for the treatment group before and after the intervention to the changes in outcomes for the control group over the same period, we can isolate the effect of the treatment. The general formula for the diff-in-diff estimator is:

$$Y = \beta_0 + \beta_1 \cdot (\text{UK}) + \beta_2 \cdot (\text{afterBrexit}) + \beta_3 \cdot (\text{UK} \times \text{afterBrexit}) + \epsilon \tag{2}$$

where:

- Y is the outcome variable (relative search volume, RSV).
- UK is a binary variable indicating whether the observation is from the UK (1 if UK, 0 if US).
- afterBrexit is a binary variable indicating whether the observation is from the post-Brexit period (1 if after June 23, 2016, 0 if before).
- UK $\times$ afterBrexit is the interaction term between being from the UK and being in the post-Brexit period.
- β_0 is the intercept, representing the baseline outcome in the US before Brexit.
- β_1 captures the difference in outcomes between the UK and the US before Brexit.
- β_2 captures the difference in outcomes in the US before and after Brexit.
- β_3 is the diff-in-diff estimator, capturing the causal effect of Brexit on the outcome in the UK relative to the US.
- ϵ is the error term.

In this study, we apply the diff-in-diff method to assess the impact of the Brexit referendum on the relative search volumes (RSV) for various keywords. We assume that the US serves as a control group that was not affected by Brexit, while the UK is the treatment group. To perform the regression and the subsequent analysis we used the Statsmodels library.

4.4 Synthetic controls

The Synthetic Control Method (SCM) is used for causal inference in comparative case studies. It constructs a synthetic version of a treatment group by optimally weighting control units (donor pool) that were ideally not affected by the treatment. A good introduction of the method can be found here. In this study, we use SCM to estimate the impact of the Brexit referendum on relative search volumes

(RSV) for various keywords.

First, we need to select a donor pool: we decided to choose large English speaking countries, namely the US, Canada, Australia and New Zealand. Second, this data is used to determine weights for control units in order to minimize the difference between the treatment unit (UK) and the weighted average of control units in the pre-treatment period. As a last step we perform a post-treatment comparison: The outcome of the treatment unit is compared with the synthetic control in the post-treatment period.

More Formally: Let Y_{it} be the outcome for unit i at time t. For the treatment unit ($i = 1$) and control units ($i = 2, \ldots, J+1$), the weights $W = (w_2, \ldots, w_{J+1})'$ are chosen to minimize:

$$\sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j Y_{jt} \right)^2 \tag{3}$$

subject to $w_j \geq 0$ for all j and $\sum_{j=2}^{J+1} w_j = 1$.

The synthetic control outcome in the post-treatment period is:

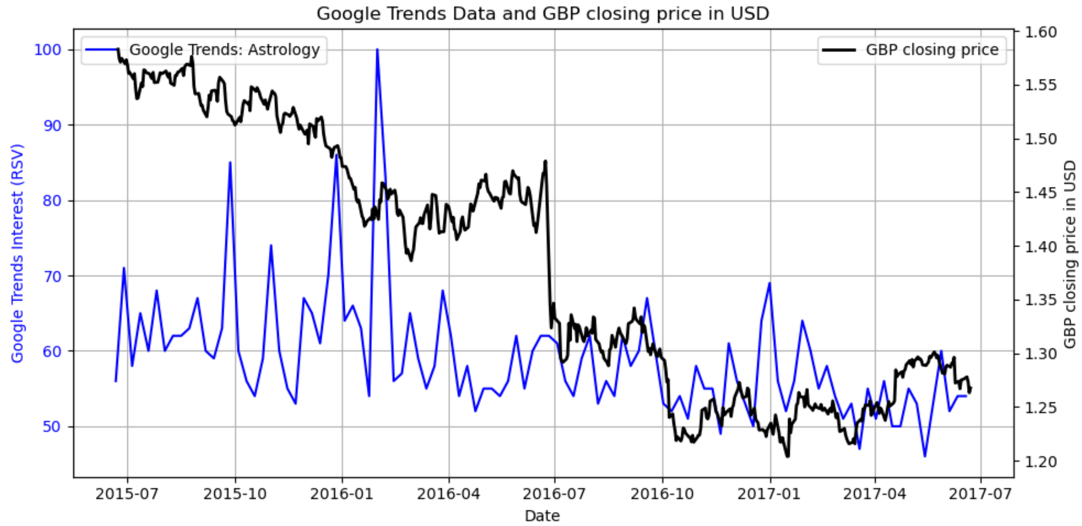
$$\hat{Y}_{1t}^{\text{SC}} = \sum_{j=2}^{J+1} w_j Y_{jt} \tag{4}$$

To evaluate the fit we use the least squares metric to assess if the synthetic control was diverging from the actual RSV for each keyword. The first step of the post treatment comparison can be done by visual inspection, where one plots the actual vs. synthetic control series and visually inspect for differences before and after the treatment period. In order to quantify the effect the root mean squared error (RMSE) can be calculated for the pre-treatment and post-treatment periods to measure the fit of the synthetic control to the actual data. A significant increase in RMSE post-treatment indicates a treatment effect.

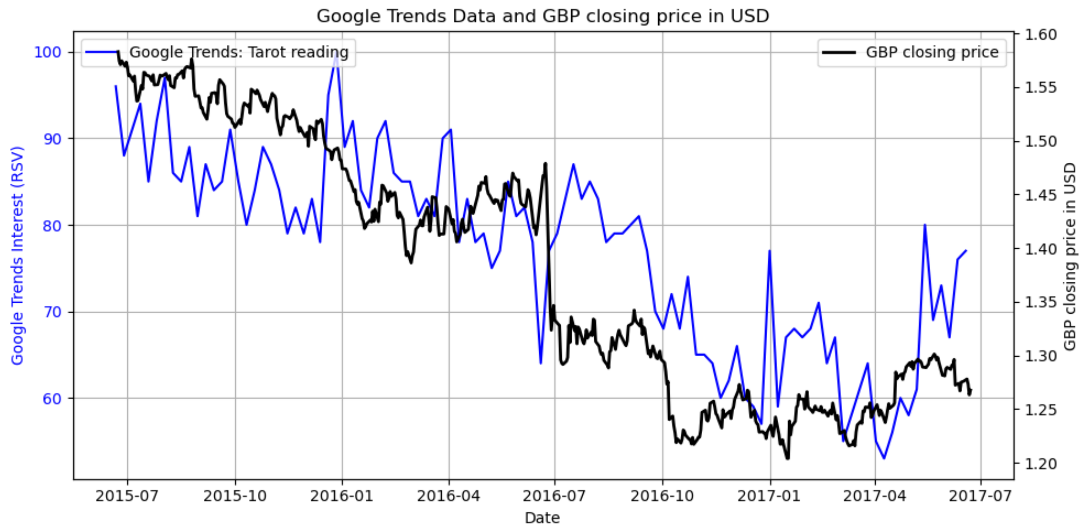
5 Results

5.1 Correlation between RSV and indicator data

To examine present correlations in the data we compare the time series of the exchange rate to the RSV of the keywords. The plot in Figure 3 shows an example of the keywords 'Astrology' and 'Tarot reading' overlayed with the exchange rate. The plotted data serves as the raw data for the calculation of the Spearman Rank Correlation Coefficient. At first sight one can notice a difference in the fluctuation of the



(a) Astrology trends



(b) Tarot reading trends

Figure 3: Overlay of trends with the closing price of GBP in USD for two example search terms. This data was used to calculate the Spearman Rank Correlation Coefficient. In Figure (a) the search term is 'Astrology' and in Figure (b) 'Tarot reading'.

data. The RSV is much more volatile than the exchange rate. Since the data are normalized between 0 and 100 and are recorded at rather large time intervals of one data point per week, the maximum value of 100 results in a relatively large peak. This can be well observed in Figure 3a. This high fluctuation prevented the calculation of the correlation coefficient from arriving at a value that actually indicates an existing correlation. The calculated values together with the corresponding P-values can be seen in Table 2. When looking at the values, it is noticeable that apart from the search term 'Tarot reading', no value is close to 1 or -1, which means that we found no evidence of an existing correlation. Due to a relatively high value of over 0.8 and a corresponding very low P-value, the RSV of the term 'Tarot reading' can certainly be considered to have a correlation with the closing price of GBP. The raw data for the calculation of this value can be seen in the image 3b. However, as this is ultimately only one of fifteen terms that indicates a correlation at all, our hypothesis cannot be proven with the help of the Spearman

Rank Correlation Coefficient applied on our data, as this is not a meaningful result. Furthermore, in this example the RSV value falls together with the closing price of the GBP although it should rise according to our hypothesis. In other words, in this case we should see a negative correlation and not a positive one.

Search term	Spearman Corr	P-value
Astrology	4.94e-01	8.30e-08
Horoscope	4.90e-01	1.16e-07
Fortune telling	2.36e-01	1.56e-02
Psychic	-2.20e-01	2.43e-02
Tarot reading	8.06e-01	3.67e-25
Numerology	-3.00e-01	1.90e-03
Palm reading	2.86e-01	3.09e-03
Clairvoyant	2.08e-01	3.35e-02
Crystal ball	1.26e-01	2.01e-01
Magic spells	3.99e-01	2.50e-05
Occult	-2.82e-01	3.55e-03
Spiritual healing	1.30e-01	1.85e-01
Witchcraft	3.00e-02	7.61e-01
Mediumship	-3.26e-02	7.41e-01
Superstition	-2.26e-01	2.02e-02

Table 2: Pearson result

5.2 Differences

5.3 Means

Figure 4 shows the resulting plots for the before/after comparison of interest. For the UK we count 4 keywords which had a net increase in interest versus 11 which had lower interest after the Brexit referendum. This can be compared to the US where 7 keywords showed higher interest versus 8 having lower interest. The difference in RSV for most keywords is very small (<7 for 23 out of 30 examples).

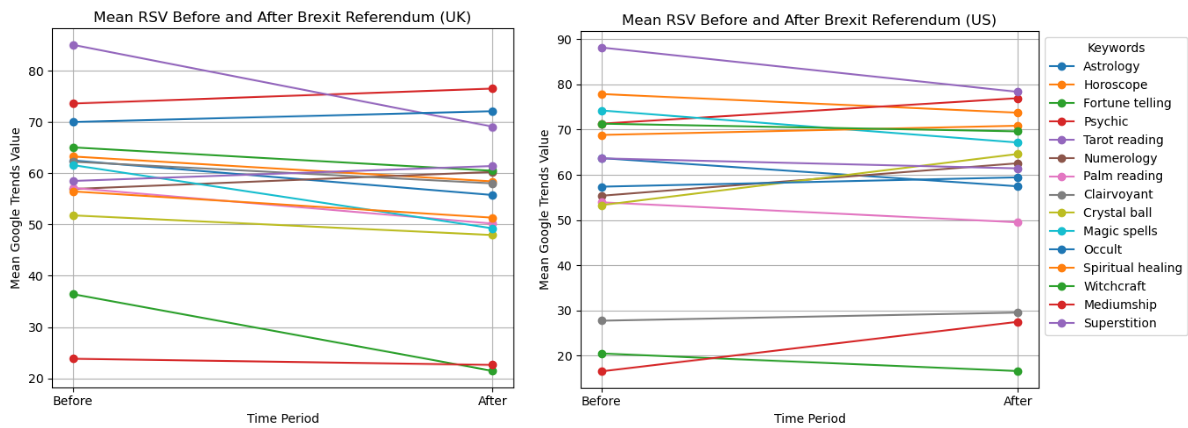


Figure 4: Comparison of mean RSV values before and after (1y each) the Brexit referendum

From this evaluation a positive trend in interest for such terms can be excluded, the data suggests that either there is no effect or it is not captured by our method. Actually, for the UK population, the total interest in magical thinking terms decreased more than for the US population. Thus if anything, interest in the observed magical thinking keywords is even tending to decline slightly after the Brexit referendum.

5.4 Diff-in-Diff

The Statsmodels library offers the framework to fit the model as well as to give a detailed evaluation which we provide in full length:

	Coef.	Std. Err.	t	P> $\|t\|$	[0.025	0.975]
Intercept	57.5862	0.756	76.197	0.000	56.104	59.068
UK	1.3635	1.069	1.276	0.202	-0.732	3.459
afterBrexit	0.0997	1.074	0.093	0.926	-2.006	2.205
UK:afterBrexit	-4.7353	1.519	-3.118	0.002	-7.713	-1.757

Table 3: Coefficient Estimates

Statistic	Value
R-squared	0.006
Adj. R-squared	0.005
F-statistic	6.771
Prob (F-statistic)	0.000151
No. Observations	3150
AIC	2.822e+04
BIC	2.824e+04
Df Residuals	3146
Df Model	3
Covariance Type	nonrobust

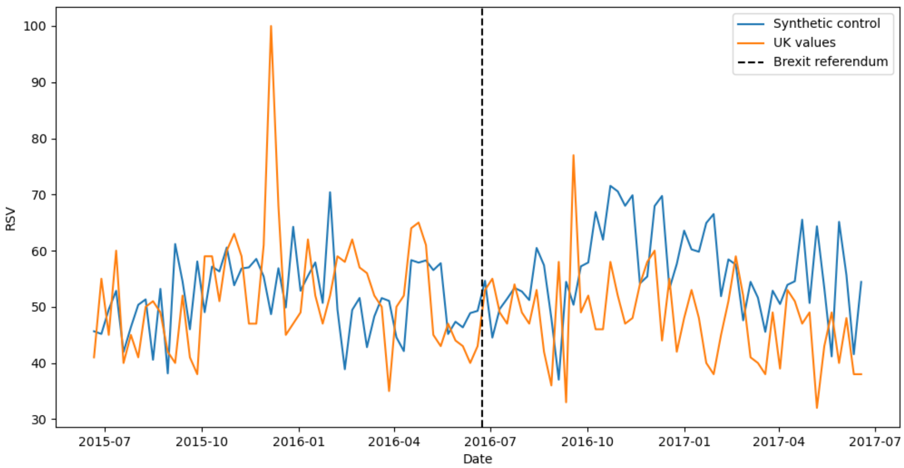
Table 4: Model Summary

The obtained results match what we could expect when having seen the previous methods. R-squared is very low which indicates that the model has very low explanatory power, this is however clearly due to the noisiness of the data. The most interesting term is the interaction term: UK:afterBrexit. It suggests that in the UK, the average RSV after the Brexit referendum decreased by 4.7353 units compared to the US.

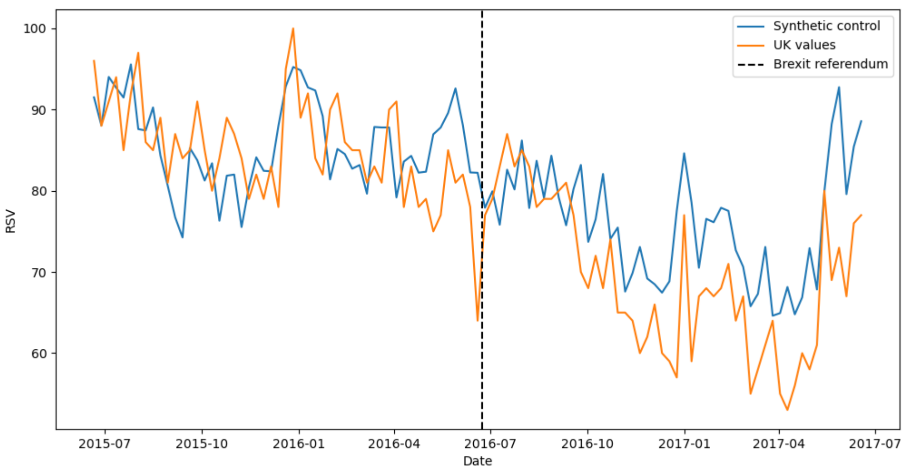
5.5 Synthetic controls

The calculated RMSE values which quantify the fit of the synthetic control to the actual data can be seen in Table 5. At first sight it becomes clear that no general significant increase of the values post-Brexit compared to pre-Brexit can be observed. This would indicate an effective treatment. What is surprising is that the values even decrease for some search terms. This means that fit is better after the Brexit even though we optimized the weights using the data only from before the Brexit. A possible explanation for this unexpected behaviour is that as already has been said several times the data looks very noisy. As can be seen in Figure 5 the graphs are fluctuating strongly due to the poor resolution of the data points. This induces a lot of randomness and makes it difficult to obtain a precise and meaningful fit. In summary, it can be said that with this method no treatment effect can be determined.

(a) Search term: 'Psychic'



(b) Search term: 'Crystal ball'



(c) Search term: 'Tarot reading'

Figure 5: Visual inspection of the actual vs. synthetic control series for three example search terms (all others can be found in the Github repository)

Search term	RMSE before Brexit	RMSE after Brexit
Astrology	10.55	9.92
Horoscope	4.07	5.38
Fortune telling	29.36	24.07
Psychic	6.95	5.80
Tarot reading	6.48	9.44
Numerology	6.80	8.09
Palm reading	11.02	14.84
Clairvoyant	12.17	11.10
Crystal ball	11.59	14.20
Magic spells	15.36	12.48
Occult	11.70	10.66
Spiritual healing	27.00	29.52
Witchcraft	11.07	13.61
Mediumship	39.18	38.35
Superstition	10.55	13.15

Table 5: Calculated RMSE values for quantifying the fit of the synthetic control to the actual data

6 Conclusion

Our project explored the relationship between economic shocks, specifically the Brexit referendum, and the interest in 'magical thinking' as reflected in Google search trends. Using a combination of Difference-in-Differences (DiD) and Synthetic Control Methods (SCM), we aimed to quantify changes in search volumes for terms associated with magical thinking in the UK compared to other English-speaking countries.

Our analysis reveals that the Brexit referendum had no statistically significant impact on the UK's interest in magical thinking. This is reflected in all of our tests. This suggests that, post-Brexit, there was no significant change response reflected in the relative search volumes for these terms in the UK. Our evaluation was not able to reject the null hypothesis (no effect of Brexit on RSV for given terms).

The study faced limitations, primarily due to the normalization, noisiness and volatility of Google Trends data. While Google Trends provides a valuable and easily accessible measure of public interest, it has several shortcomings in scientific research. The data is inherently noisy and can be influenced by various factors unrelated to genuine interest, such as media coverage or sudden social events. Additionally, relative search volumes make it difficult to compare interest levels across different regions or time periods directly. In conclusion, researchers should consider complementing or even substituting Google Trends data with other sources, such as social media analytics, survey data, or economic indicators, to achieve more accurate and comprehensive insights.

Future research could expand the scope to include other economic shocks, such as the COVID-19 pandemic, to obtain more robust findings. Additionally, incorporating more refined data collection methods or alternative sources of public interest indicators could help mitigate some of the limitations encountered in this study.

In conclusion, this research contributes to understanding how Google Trends Data should be used. It also showed a broad range of statistical tools commonly used in econometrics on a range from very simple (correlation) to more complex (synthetic controls).

Appendix

A.1 Prompt for keyword generation:

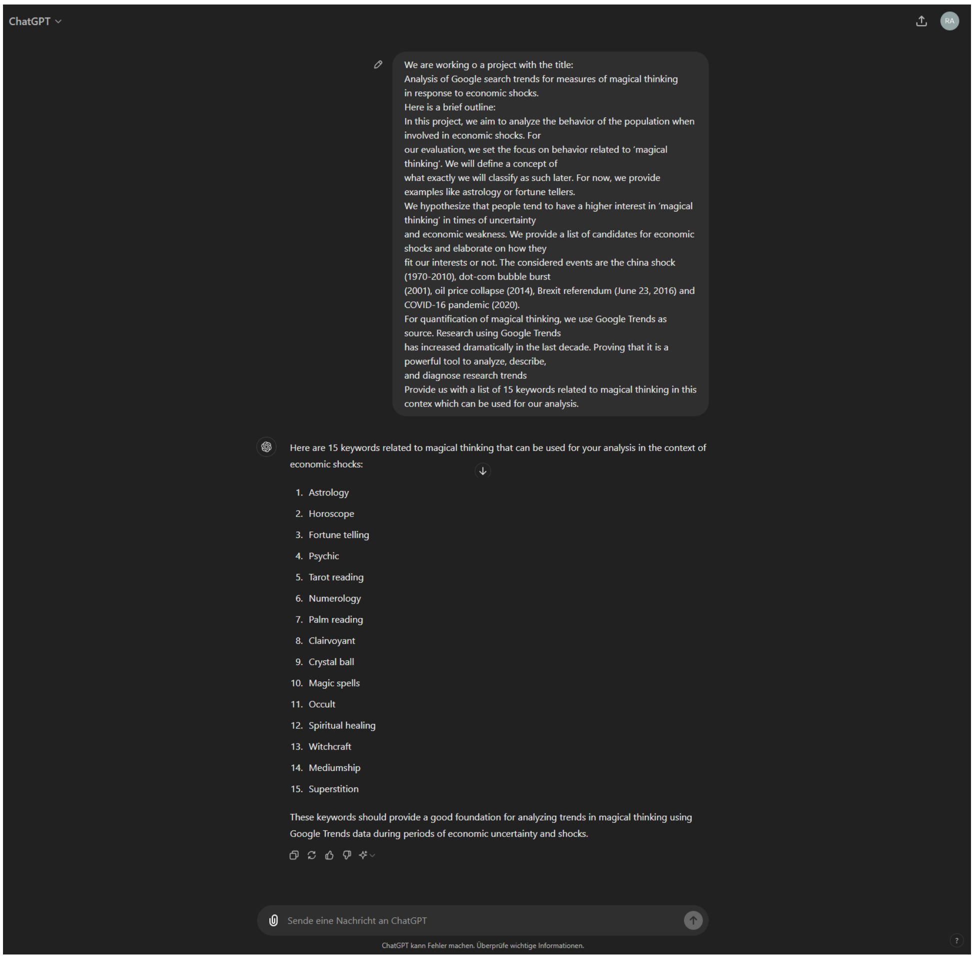


Figure 6: Caption

A.2 Replication Code

The full code as well as the data used can be found in this Github repository. Please read the README file for further information.

References

[1] Herman Anthony Carneiro and Eleftherios Mylonakis. "Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks". In: *Clinical Infectious Diseases* 49.10 (Nov. 2009), pp. 1557–1564. ISSN: 1058-4838. DOI: 10.1086/630200. eprint: https://academic.oup.com/cid/article-pdf/49/10/1557/11061074/49-10-1557.pdf. URL: https://doi.org/10.1086/630200.

[2] Tobias Preis, Helen Susannah Moat, and H Eugene Stanley. "Quantifying trading behavior in financial markets using Google Trends". In: *Scientific reports* 3.1 (2013), pp. 1–6.

[3] Leonard Zusne and Warren H Jones. *Anomalistic psychology: A study of magical thinking*. Psychology Press, 2014.

[4] Seung-Pyo Jun, Hyoung Sun Yoo, and San Choi. "Ten years of research change using Google Trends: From the perspective of big data utilizations and applications". In: *Technological Forecasting and Social Change* 130 (2018), pp. 69–87. ISSN: 0040-1625. DOI: https://doi.org/10.1016/j.techfore.2017.11.009. URL: https://www.sciencedirect.com/science/article/pii/S0040162517315536.

[5] Amaryllis Mavragani, Gabriela Ochoa, and Konstantinos P Tsagarakis. "Assessing the methods, tools, and statistical approaches in Google Trends research: systematic review". In: *Journal of Medical Internet Research* 20.11 (2018), e270.

[6] Boriss Siliverstovs and Daniel S Wochner. "Google Trends and reality: Do the proportions match?: Appraising the informational value of online search behavior: Evidence from Swiss tourism regions". In: *Journal of Economic Behavior & Organization* 145 (2018), pp. 1–23.

[7] Jadwiga Hamulka et al. "Dietary Supplements during COVID-19 Outbreak. Results of Google Trends Analysis Supported by PLifeCOVID-19 Online Studies". In: *Nutrients* 13.1 (2021). ISSN: 2072-6643. DOI: 10.3390/nu13010054. URL: https://www.mdpi.com/2072-6643/13/1/54.

[8] Google News Initiative. *Google Trends Lesson*. https://newsinitiative.withgoogle.com/de-de/resources/trainings/google-trends-lesson/. Accessed: 2024-06-10. 2024.

[9] OpenAI. *ChatGPT*. https://chatgpt.com/. Accessed: 2024-06-09. 2024.

[10] investing.com. URL: https://www.investing.com/currencies/gbp-usd-historical-data (visited on 06/07/2024).

[11] statistics.laerd.com. URL: https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-statistical-guide.php (visited on 06/09/2024).