



Contents lists available at ScienceDirect

Technological Forecasting & Social Change



YES or NO: Predicting the 2015 GReferendum results using Google Trends

Amaryllis Mavragani, Konstantinos P. Tsagarakis *

Business and Environmental Technology Economics Lab, Department of Environmental Engineering, Democritus University of Thrace, Vas. Sofias 12, Xanthi, Greece

ARTICLE INFO

Article history:

Received 11 November 2015

Received in revised form 29 April 2016

Accepted 30 April 2016

Available online xxxxx

Keywords:

Big Data

Elections

Google Trends

Internet behavior

Predictions

Referendum

ABSTRACT

We examine the possibility of predicting the 2015 Greek Referendum results by analyzing data from Google Trends on the 'YES' and 'NO' search terms. Our analysis shows that, despite the voting intention polls of the YES and NO votes being marginally one above the other throughout the prevoting period, the NO hits are clearly and every day above the YES ones, with statistically significant evidence. By analyzing data from Google Trends, we calculate a valid approximation of the final result, thus contributing to the discussion of using Google Trends as an elections' results prediction tool in the future.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Big Data are characterized by the 3 Vs: "Volume" – exponentially increasing volumes- (Hilbert and Lopez, 2011), "Variety" – wide range of datasets- and "Velocity" – high processing speed- (Chen and Zhang, 2014). Following the increased integration of Big Data in research (Gandomi and Haider, 2015), new tools of downloading, storing and analyzing data are needed, as conventional ones are not adequate (Chang et al., 2014). A recent Big Data trend is the use of Google Trends (trends.google.com), an open tool provided by Google to mainly show what was and is trending. It is rapidly becoming quite popular in academic research in various fields, such as medicine (Davidson et al., 2015), the environment (McCallum and Bury, 2013), economics and finance (Kristoufek, 2013a; Alanyali et al., 2013; Kristoufek, 2015; Choi and Varian, 2012), politics (Polykalas et al., 2013b; Huberty, 2015; Rill et al., 2014) and behavior (Preis et al., 2013; Kristoufek, 2013b).

Google Trends has been proven an effective tool in predictions, nowcastings and forecastings. Vicente et al. (2015) use Google Trends as a tool for forecasting unemployment rates in Spain, and Jun et al. (2014b) for analyzing search traffic in order to forecast sales volumes. Vosen and Schmidt (2011) highlight the performance of the indicators based on Google data and suggest that "...incorporating information from Google Trends may offer significant benefits to forecasters...". In addition, it is suggested that Google Trends can positively affect forecastings' accuracy (Han et al., 2012), and the analysis of search traffic data in general has significant potential in improving forecastings (Jun et al.,

2014a) and in analyzing online interest (Jun and Park, 2016). Finally, McCallum and Bury (2014) highlight on the validity of the Google Trends' data, and the high integration of its use in academic research.

Could we, by analyzing potential voters' internet behavior, predict elections results? As the internet's role is becoming all the more important (Wagner et al., 2016), online data could provide valuable results in measuring levels of behavioral variations (Burnap et al., 2015). Up to this point, Google Trends has been used for the development of an algorithm for predicting elections results – case study of German elections – (Polykalas et al., 2013b), and for predicting the outcome of 6 races in Greek and Spanish elections (Polykalas et al., 2013a), as this 'poll taking' method seems to be more effective. Furthermore, Reilly et al. (2012) used Google queries to measure "the public agenda" in the 2008 US Presidential elections. Subject to careful selection of the searched terms, Google data can accurately measure the public's interest (Scharkow and Vogelgesang, 2011). Big Data brings political science to a whole new level, due to the increased internet penetration and the use of online tools and social media by political campaigners (Weber et al., 2013).

The political situation in Greece has been unstable over the course of the last few years, with 4 election races and one referendum in less than 3.5 years. This has led to an increased online interest in politics, with high percentages of internet users becoming active in social media in political issues, resulting in the potential voters' notable changing in internet behavior.

On Saturday, June 27th 2015, Greece's Prime Minister makes a live address to the nation, abruptly announcing that a referendum is to take place, just a week in advance on Sunday, July 5th (Yardley and Kitsantonis, 2015). This resulted in the confusion amongst the voting population and the politicians, due to the short 1-week prevoting period and the not so clear question that the voters were asked to answer to. A

* Corresponding author.

E-mail addresses: amavraga@env.duth.gr (A. Mavragani), ktsagar@env.duth.gr (K.P. Tsagarakis).

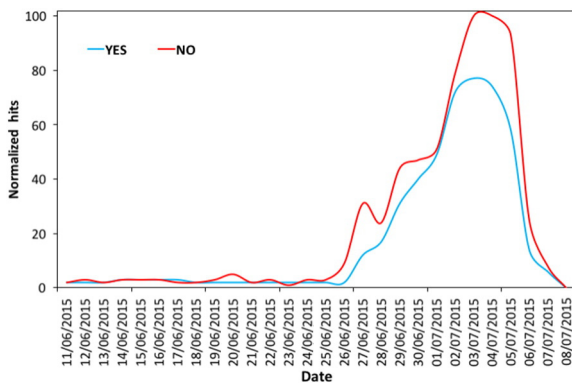


Fig. 1. YES and NO hits in Google from 06.11.2015 to 07.07.2015.

“NO” response could result in Greece’s leaving the Eurozone or the European Union (European Commission, 2015) –followed by any potential consequences to Greece’s economic situation. The 2015 GReferendum received wide publicity in national and international media, with mass writing, debates and live news updates and coverages on the subject emerging over the prevoting period (BBC, 2015).

Our aim is to look into the possibility of predicting the 2015 Greek Referendum voting intentions and results by analyzing data from Google Trends. The Referendum was announced on June 27th, 2015 and took place on July 5th, 2015. The voting population was asked to respond to a dilemma concerning a European rescue plan with a “NAI” or “OXI” (meaning YES and NO, respectively) (Yardley and Kitsantonis, 2015). Based on the idea of Preis et al. (2012), who use only Arabic numerals as keyword Google searches, we view the 2015 Greek Referendum as a unique opportunity to examine this possibility, as the only options are YES and NO, and no linguistic differences, misspellings or translation errors could provide invalid results. The rest of the paper is structured as follows: in Section 2 the data and the research methodology are presented, followed by our results and discussion in Section 3, and Section 4 consists of the concluding remarks.

2. Data and method

The Google search activity in the days following the Referendum announcement was rising, and we noticed that the daily hits for ‘NO’ were clearly and consistently higher than the ‘YES’ ones in Google Trends. Official voting intention polls, on the contrary, suggested that it would be a close battle, in some cases giving as little as a 0.5% difference between the ‘YES’ and the ‘NO’ vote (Sembhy, 2015).

As shown in Fig. 1, the internet activity on the YES and NO searches suddenly increased on the 27th of June, just after the Referendum announcement, and continued rising the following days. The curve peaks

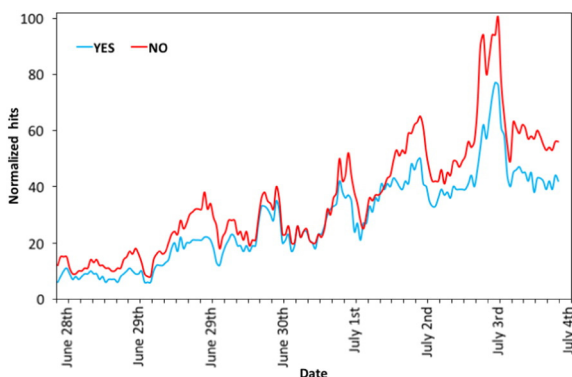


Fig. 2. YES and NO hits in Google from June 27th (20:00) to July 4th (20:00).

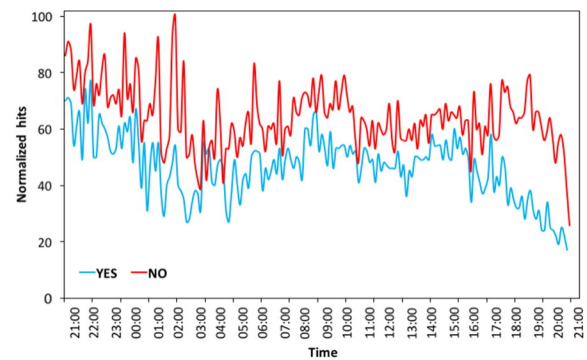


Fig. 3. YES and NO hits in Google from Saturday the 4th (20:00) to Sunday the 5th (20:00).

on July 5th, the day of the Referendum race, and starts decreasing afterwards, reaching again the normal searched-for levels. This shows that the YES and NO searches are linked to the referendum, as the search volumes on the relative week show about 98% increase than the ones before the race (also see Supplementary Material Online, Figs. S2 and S3).

Fig. 2 shows the graph of the normalized data from Google Trends for the YES and NO hits during the prevoting period, i.e. Saturday, June 27th to Saturday, July 4th. Once again, we see that the NO hits are consistently above the YES ones, supporting our hypothesis that the NO Vote seemed to be leading the July 5th Referendum race.

Figs. 3 and 4 show, respectively, the graphs of the normalized data for the YES and NO hits from Saturday the 4th (20:00) to Sunday the 5th (20:00), and on the referendum day, July 5th, from 7:00 to 19:00, opening and closing of the ballot boxes, respectively. Based on the graphs, and seeing that the Google Trends’ data paint a different picture than the official published voting intentions, we percentize the normalized data for each set period in order to predict the referendum outcome.

Data provided by Google Trends are already normalized over each selected period and downloaded online in .csv format. The normalization process is reported as follows: “each data point is divided by the total searches of the geography and time range it represents, to compare relative popularity. The resulting numbers are then scaled to a range of 0 to 100” (Google Trends, 2016).

Starting on Saturday, June 27th, we download data from Google Trends. The data we use in our analysis are hourly intervals (daily data) of the period Saturday, June 27th (21:00), to Saturday, July 4th (20:00) – downloaded at 23:35 each day – (hereafter Set1), 8 min intervals (4-hour data) from Saturday, July 4th (21:00), to Sunday, July 5th (20:00) – downloaded at xx:35 every 4 h – (hereafter Set2), and 1 min intervals (hourly data) on Sunday July 5th from 07:00 to 19:00 – downloaded at xx:35 of each hour – (opening and closing of the ballot

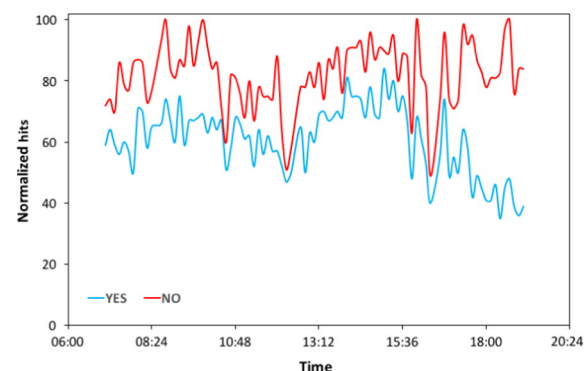


Fig. 4. YES and NO hits in Google on Sunday the 5th from 7 am to 7 pm (8-min intervals).

Download English Version:

<https://daneshyari.com/en/article/7256012>

Download Persian Version:

<https://daneshyari.com/article/7256012>

[Daneshyari.com](https://daneshyari.com)