IMPACT OF BIG DATA AND POLITICAL MICROTARGETING ON DONALD TRUMP’S 2016 PRESIDENTIAL CAMPAIGN

Narishia Singh1 and Anuradha Wijegunawardhana2
1nsingh@nrcocp.com, USA
2anuradhawije@nrcocp.com, Sri Lanka

Abstract- On January 20th, 2017, Donald Trump was inaugurated as the 45th President of the United States. His opponent, Hillary Clinton, was a seasoned politician who was favored to win the election. However, Trump produced one of the biggest political upsets in modern day history. This explanatory case study aimed to explain one of the reasons that Trump won by looking at the role that technology played during the 2016 presidential election. Specifically, this study examined how big data was leveraged to influence voter behavior during the campaign. Data drawn from articles, interviews, and videos published during the campaign were collected and analyzed thematically. Results identified three main themes that were particularly salient: mobilization of voter databases, data-driven microtargeting, and Facebook influence. This study highlights the potential benefits of targeting voters using big data analytics as well as the potential risks related to issues of privacy.

Keywords- Big Data, political microtargeting, psychographics, election campaigns

I. INTRODUCTION

The proliferation of technology in society has had a major impact on political campaigns. The field of big data specifically played a crucial role in the recent 2016 US presidential election with the rise of data-driven political microtargeting fuelled by the Trump campaign. Some analysts note that Trump’s victory was based on his campaign’s ability to target specific segments of the electorate more effectively than Clinton. This explanatory case study thereby aimed to understand how big data and political microtargeting were used during Trump’s 2016 presidential campaign. The following question guided the study: 1) How did the Trump campaign leverage big data to influence voter behaviour?

Prior research that examined how political candidates in the U.S. gather and manipulate data to appeal to voters, has focused on the Obama campaigns mainly (Tufekci, 2014). This case study will, therefore, contribute to research in this field.

II. LITERATURE REVIEW

Big data election campaigns have gained increased momentum in the last decade. Access to a huge volume and variety of voter data, along with a growing infrastructure of cutting-edge technology tools now allow campaign teams to gain deep insight about each constituent. Although the term itself has numerous definitions, for the context of this paper, big data is defined as “large pools of data that can be captured, communicated, aggregated, stored, and analyzed” (McKinsey Global Institute, 2011, p. 1). Analytics is a key aspect of big data since it allows data-driven decisions to be made. “To help decision making, data analysts choose informative metrics that can be computed from available data with the necessary algorithms and tools and report the results in a way the decision makers can comprehend and act upon” (Lau, Yang-Turner and Karacapilidis, 2014, p. 50). Candidates now rely heavily on data collection and analytics to shape their communication strategies, specifically as it relates to political microtargeting. “Through political microtargeting, a political party can identify the individual voters which it is most likely to convince” (Borgesius et al., 2018, p. 82). They can then use various campaign appeals (e.g. door-to-door conversations, direct mail, social media messages, emails or texts) to connect with these voters using persuasive messages.

Research about big data and political microtargeting can be grouped into three broad topical categories: voter data, political data analytics, and voter microtargeting. A brief overview of each category is discussed below.

1. Voter Data: The United States has maintained electronic voter profiles with data points for each registered voter, for over two decades. Research by Alexander and Mills (2004) show that state voter registration databases usually contain the name, address, date of birth, and phone number of each individual along with past electoral participation. On a national level, both the Republicans and Democratic parties maintain expansive digital databases (known as voter files) that merge the states’ voter data with other data forms of public data (e.g. motor vehicle, real estate and campaign donation records) and consumer data (e.g. credit reports). As a result, a voter file can be acquired from data brokers (Howard and Kreiss, 2010; Rubinstein, 2013). For example, as of 2015 the Republican party’s GOP Data Center contained information on over 200 million individuals with over 700 billion data points” (GOP, 2015). An increasing number of partisan and nonpartisan political firms also offer commercial voter files. As an example, long-standing non-partisan firm Aristotle maintains a national voter file with “over 190 million records that are enhanced with phones, emails, demographics and lifestyle information” as well as a separate consumer database and donation file (Aristotle). Akosah (2015) notes that the ongoing rise of social media and digital devices means that voter files are constantly updated with new information from online sources, including campaign website registration and online tracking cookies. Voter data is also accumulated from offline sources such as volunteer and fundraising events. (Akosah, 2015).

2. Political data analytics: Campaigns analyze voter data “to form predictions about which members of the public are supporters, which are likely to show up to vote, and which are persuadable” (Hersh, 2015, p. 3). Based on these predictions, campaigns can make strategic decisions about who to target and formulate fundraising objectives. Nickerson (2014, p.58) notes that by using voter data, “campaigns are able to predict with greater accuracy which citizens will support their candidature and issues better than which citizens will oppose [them].” For example, by looking at the attributes of citizens who donate or volunteer, campaigns can design models to identify correlations among similar citizens (Nickerson,2014). Data analysts use sophisticated statistical techniques to create predictive models from a universe of voters; such as a contactable, support, turnout and persuasion model.

For example, during his re-election campaign in 2012, Barack Obama’s campaign created a “likelihood to turnout” model using surveys and the vast datasets collected about each voter (Tufekci, 2014). After assigning a score from 0-100 to determine a voter’s likelihood of voting, the campaign used this information to target individuals who scored high on the index (Tufekci, 2014).

3. Voter Microtargeting: Vast sets of voter data allow campaigns to target individual constituents with specific messages. For example, instead of targeting broad demographic communities (e.g. women or minorities), campaigns can segment the population and appeal to a conservative white male living in Jacksonville, Florida or a liberal African-American living in Philadelphia. According to Howard and Erickson (2009), this benefits a campaign since they can allocate valuable funds towards voters who are receptive to their message rather than focusing on unlikely voters. Microtargeting works by first collecting all of the aggregated information in each voter’s profile including survey, “demographic, geographic and marketing data…” (Bennett, 2015, p.374). Predictive models are then created and “hundreds of customized messages for each [constituent], each with highly personalized political content...” are designed and then delivered using various channels of communication (International Institute for Democracy and Electoral Assistance, 2018, pp.13-14). Recent research conducted by Schipper and Woo (2017, pp. 18, 22) shows that microtargeting is necessary for a voter to view a candidate favorably since individuals who do not

Figure 1. Creating Persuasion Universes in Battleground States Source: Fulgoni, Lipsman and Davidson, 2016
III. METHODOLOGY

This study followed a qualitative method of research using an explanatory, single-case study. In defining what a case study is, Yin suggests that the term refers to an event, an entity, an individual or even a unit of analysis. It is an empirical inquiry that investigates a contemporary phenomenon within its real-life context using multiple sources of evidence (Yin, 1989). An explanatory case study was selected as the preferred strategy since it specifically aims to answer “how” or “why” questions in-depth, using multiple sources.

Data collection consisted of an eleven-minute online video of Alexander Nix, the former CEO of Cambridge Analytica who gave an overview of big data and psychographics and highlighted the company’s collaboration with Senator Ted Cruz’s presidential campaign. Similarly, a 60 minutes interview featuring Trump’s digital director Brad Parscale was included. In addition, an online article published by the magazine Businessweek was also sourced since it focused on the Trump campaign’s real-time data operation twelve days before the presidential election. Two-thirty-second advertisements that the campaign aired were also included. To add further insight about this research, a dataset of 24 online documents related to big data and political microtargeting was sourced using the Google Scholar search engine. To address the issue of validity and reliability, triangulation of the data was entailed; documents were drawn from various disciplines including communication, economics, law, marketing and political science.

In terms of analytical strategies, thematic analysis was selected since it provides a systematic yet flexible element to data analysis. As Braun and Clarke (2013, p. 6) note, “thematic analysis is essentially a method for identifying and analyzing patterns in qualitative data.” The goal is to identify themes from a broad set of data by determining the relationships between concepts. Braun and Clarke’s iterative six-step framework is widely considered the most influential approach for doing thematic analysis and was therefore utilized in this study.

IV. RESULTS

Results show that there were three ways the Trump campaign leveraged big data to effectively target key voters.

1). Mobilization of Voter Databases: The Trump campaign accessed a trove of voter data by utilizing three voter databases during the election. First, was the deployment of a custom-designed fundraising database named Project Alamo which was initiated in November 2015. The database quickly expanded after Trump became the Republican nominee for President in July 2016. Data was gathered from small donor contributions, email addresses gathered at rallies, sales of campaign merchandise and text messages sent to the campaign (Halpern, 2017). Likewise, Cambridge Analytica also designed a specialized database for the campaign; however, it was based largely on psychographic information about America voters. In a presentation given at an annual non-partisan summit in September 2016, Alexander Nix described how his political consulting firm administered a Big Five personality survey (they gave it the acronym OCEAN) to hundreds and hundreds of thousands of Americans (Concordia, 2016). Along with the survey, the company also acquired data about individuals using sources such as “…digital data, voter history, and marketing resources supplied by leading companies, including Acxiom, Expertian, Nielsen…and Facebook…” (Chester and Montgomery, 2017). Cambridge Analytica claimed that all this data was then aggregated using big data analytics to create a personality profile for every adult in the United States.

Finally, the campaign relied on the GOP Data Center, the 20-year old national voter file. Eventually, Cambridge Analytica’s database and the GOP Data Center would be merged with Project Alamo, to create a massive database that held the profiles of over 220 Americans. Each profile “contained between 4,000 and 5,000 data points...of individual voters” (Owen, 2018, p. 41).

2). Data-Driven Microtargeting: Using the data gathered on the electorate, the Trump campaign micro-targeted precise segments of the population. For example, data from the GOP Data Center revealed that infrastructure was an issue of concern to some voters. In a 2017 profile on 60 minutes, the campaign’s digital director Brad Parscale discussed how “voters in the Rust Belt cared about their roads being built, their highways, their bridges” (CBS News). He then started making ads that would address these issues to target persuadable voters in states such as Pennsylvania and Michigan. The campaign continued to focus on the Rust Belt states as the campaign progressed.

Figure 2. Same demographics, different personalities
Source: Computer Business Review
In October 2016, the campaign’s data analytics team (selected Facebook users (Green and Issenberg, 2016). In his interview, Parscale claimed that “he calculated early that Facebook was how Donald Trump was going to win” and he spent much of his budget on Facebook ads (CBS news). For example, in the early stages of the primaries, Parscale launched microtargeted advertisements on Facebook using data collected about all known Trump supporters. The advertisements relied heavily on Facebook’s powerful custom audience targeting features which allowed the campaign to target voters based on interests and behaviors (Halpern, 2017). From there, Parscale expanded his pool of targets using “…Facebook’s Lookalike Audiences tool to find people with interests and qualities similar to those of his original cohort and developed ads based on these characteristics…” (Halpern, 2017). During the third presidential debate, the campaign also tested 175,000 different variants of an advertising post on Facebook. “The messages differed for the most part only in microscopic details, in order to target the recipients in the optimal psychological way: different headings, colors, captions, with a photo or video” (Grassegger and Krogerus, 2017).  

V. DISCUSSION

This study examined how big data and political microtargeting were used during the 2016 US presidential election campaign. As the results show, the widespread availability of huge sets of voter data made it possible for the campaign to gain deep insight about each voter during the election. The campaign then used this information to create ads that targeted specific individuals. The use of psychographics added an advantage; as Alexander Nix, the former CEO of Cambridge Analytica noted, “If you know the personality of the people you are targeting, you can nuance your messaging to resonate more effectively with those key audience groups” (Concordia, 2016). Finally, Facebook offered a broad spectrum of commercial digital marketing tools and techniques which the campaign used to appeal to voters.

From a broader perspective, several key observations can be made from the research. First, Trump’s 2016 campaign shows that big data now plays a critical role in politics. Campaigns have access to a vast amount of voter data which they can analyze to create a comprehensive profile of each member of the electorate. Elections are also becoming more localized; as campaigns shift away from targeting voters across broad communities. At the same time, social media continues to be a main communication channel between candidates and voters. Nevertheless, this study raises several important issues about data privacy. Concerns have been raised about the way Cambridge Analytica accessed voter data and the company dissolved earlier this year as media reports indicated that the company allegedly received psychographic data from millions of Facebook users without their consent.

Questions about ethics also arise because of the Cambridge Analytica scandal. Big data allows for highly personal information to be mined about people on a massive scale, yet privacy regulation in the United States is very weak. Individuals, therefore, have no significant data protection when their data is breached or used without consent. Having access to people’s digital footprints also creates a threat of manipulation with political microtargeting. "For instance, a party could target particular voters with tailored information that maximizes, or minimizes, voter engagement" (Borgesius et al., 2018). One, therefore, must wonder whether politicians will use the valuable insight that big data provides to make unfair or discriminatory decisions during an election campaign.

VI. CONCLUSION

The Trump presidential campaign of 2016 revealed how big data and political microtargeting are transforming political elections. Through an analysis of data sources, this case study revealed that big data was used to tactically target specific individuals and make large-scale strategic decisions. From a broader perspective, this study demonstrates how a mass analysis of individual data can have a strong impact on both political communication and the political system itself. Nevertheless, this study has several limitations. Campaigns driven by big data have a new phenomenon and although important journalistic commentary exists about the topic, academic research is fragmented. For example, there is very little research available about the methodology of mining voter data to discover correlations and patterns. Likewise, as a single case study, this research cannot be generalized to a wider population. Further research that investigates how big data and political microtargeting are used in elections will, therefore, be fruitful.

REFERENCES


Binary logistic regression, Data mining, empirical study has proved the effectiveness of our proposed approach. The disease reveal as an autoimmune disease of diabetes patients experience this type of Diabetes. Though there are actually only around 10% type accounts for nearly 90% of the diabetes cases happening at a very young age of less than twenty years thus further called juvenile-onset diabetes. It often known as blood sugar, is simply too high. Blood sugar levels and this can lead to damage to the tissues and failure of the organs (Kelly, 2006). Any change within the food into the body cells. Any change within the that directs the glucose which is produced by breaking down the food. The body cells dismantle the food into glucose and this glucose should be shifted to all of the cells of the body. Insulin, a hormone produced by the pancreas, helps sugar with the application of the binary logistic regression model. Statistical model has been created using Minitab for the prediction of diabetes which based on 12 risk factors and computational politics. But, the amount of knowledge obtained from such data is extremely challenging task because of the complicated nature of diabetes. Early prediction of diabetes is an extremely important task.