Correlation between the Presence of Big Box Retailers and a County's Political Leanings

Pouya Mohammadi

Introduction

In recent election cycles, election analysts such as Dave Wasserman have pointed to an interesting metric and its being highly correlated with the results of United States (US) presidential elections. This metric is whether the county contains a Whole Foods or a Cracker Barrel (Wasserman, 2020). This has lead to a flurry of analyses about how the presence of different big box retailers is correlated to that area's political leanings. NBC News, Time Magazine, and The New York Times have all recently reported on similar trends with varying reatilers, and on Twitter after the 2020 election, there was speculation that the presence of a Trader Joe's was becoming a valid indicator for the direction in which a county would vote in the election ("Broke: Joe", 2021).

These trends are important as computational politics becomes of ever-increasing importance in modern political campaigns. In elections that are decided by the slimmest of margins, such as the US presidential election of 2000 that was decided by less than 600 votes, every informational advantage that a campaign has can be of use and drastically influence the politics and future of the United States and the world (Glass, 2018). By understanding the relationships between the counties that certain big box retailers are located in and the political preferences of that county, campaigns may be able to draw insights into the ways that communities' shopping habits correspond with their politics and more effectively and efficiently communicate with potential voters. In addition to providing insight into the politics of communities, this method of using the presence of big box retailers in a county to infer information about politics is computationally efficient and far less expensive than big data methods that rely on storing and analyzing information on an individual basis.

Our case study aims to analyze the presence of a new set of big box retailers in US counties and understand the ways that the presence of these counties corresponds with that county's political leanings. In particular, the stores that we will be focusing on are Trader Joe's, Cinemark, Cabela's, and Nordstrom. These stores were chosen because of our belief that they will capture information about the different voting blocks in the US electorate incredibly well. Aside from Trader Joe's, all of these stores have yet to be the subject of a rigorous analysis that relates their presence to political leanings. Trader Joe's was included because we believe that it corresponds to a different voting block than the other stores that we chose and the fact that the body of work on the political leanings of "Trader Joe's counties" is new and does not include a methodology similar to ours. Our null hypothesis is that we do not expect there to be a relation between any of the stores and the political leanings of the counties in which these retailers are located.

Data

We use publicly available data online that can be downloaded for the list of sites in the Data subsection of our References. These include open-source government databases as well as public location listings by private companies. The 2020 presidential election results that we use to measure a county's political leanings are borrowed from the GitHub user @tonmcg, who developed the datasets from information provided by The Guardian, townhall.com, Fox News, Politico, and the New York Times (Tonmcg, 2020). All of the government provided databases are estimations of the current values based onthe most US Census which occurred in 2010 and trends in the data since that time. Finally, the location listings are either provided by the stores

themselves or reputable third-party sources, such as Fandango. We use the 'geopy' Python library to map these locations to their corresponding geographical county. Any lcations whose counties cannot be found by the Python library are hand-encoded. Because Alaska does not use counties in their voting procedures, we map all of Alaska's voting districts to Anchorage County.

We will use the 2020 presidential election results from a county as our dependent variable, which will be a binary variable of whether the county voted Democratic or Republican, to measure the political leanings of a county. For our independent variables, we will have four variables that each correspond to whether a county contains one of the four retailers that are the focus of this analysis. In addition to these variables, we will include indpendent variables that are usually included in analyses of a county's political leanings. These variables are borrowed from Kahane, and they include population totals by race, population totals by gender, education levels, poverty rates, the median income of the county, unemployment rates, and whether the county is urban or rural (Kahane, 2020). We exclude certain variables that we did not have access to, such as religion in our analysis.

Existing Literature

As mentioned previously, plenty of news networks have completed analyses about the ways that big box retailers' locations correspond to political leanings in a county. In particular, NBC News, Time Magazine, and The New York Times have conduced analyses on this phenomenon. The New York Times in their analysis, found that, without controlling for other factors, the presence of a Trader Joe's corresponded with better results for Democrats. The analysis discovered that Democrats won areas that were within five miles of a Trader Joe's by 33 points. Time magazine discovered a similar trend in that Democrats won districts with a Trader Joe's by 30 points in 2014 US congressional races. A different study by Aaron Lee that employs a random forest machine learning model that uses 20 big box retailers discovered that counties that contain a Trader Joe's lean democratic with a feature importance of about 0.11 (Lee, 2020).

Despite all of the attention to these trends, there does not seem to a published work on the topic. We expand on these past analyses by analyzing a new set of big box retailers and their presence's relation to political leanings in a county. We also control for other variables that are customarily used in predicting a county's political leaning, which has yet to be done in a scientific analysis. We gather this list of variables from Kahane, and we explore whether the presence of these big box retailers will provide more significant relationships to a county's political leaning than the standard covariates.

EDA

References

Glass, A. (2018, December 12). Bush declared Electoral victor Over Gore, Dec. 12, 2000. Retrieved March 03, 2021, from https://www.politico.com/story/2018/12/12/scotus-declares-bush-electoral-victor-dec-12-2000-1054202

Kahane, L. H. (2020). Determinants of County-Level voting patterns in the 2012 and 2016 presidential elections. Applied Economics, 52(33), 3574-3587. doi:10.1080/00036846.2020.1713985

Lee, A. (2020, September 29). Trader Joe's Democrats and Walmart Republicans. Retrieved March 04, 2021, from https://towardsdatascience.com/are-you-a-trader-joes-democrat-or-a-walmart-republican-a7b156131435

M. (2021, February 16). Broke: Joe Biden did so well in counties with a Trader Joe's because the audience for Trader Joe's is composed of favorable Democratic demographics Woke: Joe Biden did so well in counties with a Trader Joe's because voters thought he was Trader Joe. Retrieved March 03, 2021, from https://twitter.com/maxtmcc/status/1361504297477890050

Wasserman, D. (2020, December 08). Fact: Biden won the presidency Winning 85% of counties with a Whole foods and 32% of counties with a Cracker barrel - the widest gap ever. Retrieved March 03, 2021, from https://twitter.com/Redistrict/status/1336342894630858755

Data

https://www.traderjoes.com/pdf/Trader-Joes-Stores.pdf http://nordstromsupplier.com/Content/sc_manual/Store_Address_List.pdf https://www.fandango.com/movie-theaters/cinemark https://stores.cabelas.com/ https://www.census.gov/data-tools/demo/saipe/#/?map_geoSelector=mhi_c&s_measures=mhi_snc&s_year=2019 https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/

Tonmcg. (2020). Us_county_level_election_results_08-20. Retrieved March 03, 2021, from https://github.com/tonmcg/US_County_Level_Election_Results_08-20

Data Modelling