

Adarsh N. Parthasarathy

Advisor: Professor Thad Kousser

Senior Thesis, Department of Political Science

30 March 2020

Talking Points: How Primary Elections Affect Expressed Political Ideology on Twitter in
U.S. Senate Elections from 2014-2018

Abstract

For a concept that has been a central tenant of political science and game theory for the better part of over 60 years, the median voter theorem has generated an unusual amount of study to both bolster its claims and defeat it once and for all. It has been blamed for polarization and heralded as the gold-standard on which politicians can ride to an electoral victory, but it has never been studied in light of modern advancements in political communication. In particular, Twitter presents a unique medium to observe whether or not candidates subscribe to the median voter theorem within the framework of the primary-general election system in the United States by observing their expressed political ideology over the course of an electoral season. This study shows that candidates for Senate exhibit an interesting ideological trend during an election that shows them appeal to their party's hypothetical median voter during the primary and pivot to more centrist positions as the general election nears. Although results for Republicans are inconclusive as to whether the openness of a primary affects expressed sentiment, Democrats show a significantly greater shift from primary to general election ideology in states with closed versus open primaries. This has important implications for future work in this realm.

Acknowledgements

I would like to acknowledge the leadership and support that my advisor, Professor Thad Kousser, has provided me across the 6 long months of writing and the 8 prior months of supporting research that helped to set me on the path for this topic. More than any other experience at UC San Diego, including my coursework, this is the experience I will look back on as what defined my time as a student here, and I would not have it without him.

I would also like to acknowledge Professor Philip Roeder and the invaluable experience he offered me. Although he did not advise me on this thesis, research work I did with him set the foundation for some of the most critical skills I used in assembling this thesis.

Thank you.

Contents

Introduction	7
Chapter 1	13
Relevance	13
History, Context, and Motivation	14
Literature Review	17
Theoretical Framework of the Median Voter Theorem in Elections	17
The Median Voter Theorem and Primaries, Reworked	24
Why Use Twitter?	26
Hypotheses	33
Chapter 2	38
Methods	38
Data Inspection	41
Chapter 3	48
Results	48
H1	49
H2	55
Summary of Results	59
Chapter 4	62
References	66

List of Figures

1	MVT example using the 2016 Presidential Election ¹	19
2	MVT with primary district-level preference overlaid. It is worth noting that party preferences would see distributions with medians closer to the periphery ²	19
3	Diagram of Supervised Machine Learning	38
4	Example of Tweet Cleaning	41
5	Mean Ideology Per Week By Party	43
6	Difference in Average Ideology Between Republicans and Democrats By Week	43

1. Jonathan Gruber, *Public Finance and Public Policy (3rd)* (Worth Publishers, 2011), 860, ISBN: 9781429219495.

2. Craig B McLaren, *DOES GERRYMANDERING VIOLATE THE FOURTEENTH AMENDMENT?: INSIGHT FROM THE MEDIAN VOTER THEOREM*, technical report (2017).

List of Tables

1	Start and End Dates for Data Collection By Election Year	40
2	Summary Statistics of the Final Data Set	44
3	Sample Tweets	44
4	Summary Statistics of Ideology by Party & Election	45
5	Primary Types and Their Respective States	47
6	Sample Tweets and Respective Coding	52
7	Linear Regression Results for H1	53
8	Fixed Effects Regression Results for H1	54
9	Linear Regression Results for H2	57
10	Linear Regression Results With Controls for H2	57
11	Fixed- and Random-Effects Regression Results for H2	58
12	Fixed- and Random-Effects Regression Results With Controls for H2	58

Introduction

The 2016 Election in the United States was, in many ways, a lesson in political extremism. After the surprise ascension of Donald Trump to the Presidency, the journalistic and academic community scrambled to try and understand how, in violation of the Downsian spatial electoral models we hold so dear, such an apparently extreme candidate would manage to capture enough of the vote to win the White House. The apparent violation of traditional ideas of what makes an electable candidate shocked the public and led to scholarship that focused on whether Trump's presentation matched what the voters thought of him and where he was positioned on an objective political scale.

As it turns out, based on the voters themselves, Trump ran as more of a moderate than Hillary Clinton. In 2016, the Pew Research Center released an online report entitled "Voters' Perceptions of the Candidates: Traits, Ideology and Impact on Issues," focusing on voter evaluations of Clinton and Trump that confirms this analysis. According to the report, 58% of voters saw Clinton as liberal on all or almost all issues and 28% saw her as a mix between conservative and liberal; Donald Trump, in comparison, saw 44% and 40% for the same questions, respectively.³ This does not paint a picture of an extremist, but rather a political moderate. As Matthew Yglesias wrote for Vox, "[Trump] paired extremely offensive rhetoric on racial issues with positioning on key economic policy topics," creating a public image that was often at odds with his expressed policy positions during the election.⁴ Trump also appears to be rated as a comparative moderate when pitted against

3. Pew Research Center, *Voters' Perceptions of the Candidates: Traits, Ideology and Impact on Issues* / Pew Research Center, 2016, accessed December 7, 2019, <https://www.people-press.org/2016/07/14/voters-perceptions-of-the-candidates-traits-ideology-and-impact-on-issues/>.

4. Matthew Yglesias, *Moderates win elections: ideological extremism is risky* - Vox, 2019, accessed December 7, 2019, <https://www.vox.com/2019/7/2/20677656/donald-trump-moderate-extremism-penalty>.

past Republican candidates for president. Gallup polls indicate that only 47% see him as a total conservative and 19% see him as a total liberal, a low and high, respectively, among all Republican nominees since George H. W. Bush in 1992.⁵ This realization should not be surprising for those who subscribe to traditional Downsian theory, which purports that "a majority rule voting system will select the outcome most preferred by the median voter."⁶

Nevertheless, Donald Trump's positioning leads to a related question regarding the apparent increase of political polarization in the United States, a condition that makes no sense in light of traditional Downsian theory. As Professor Larry Bartels from Vanderbilt University wrote in *The New York Times*, "contemporary American political parties seem not to have gotten the message" concerning the power of moderates in an election as the parties diverge from the political center more and more each year.⁷ The United States electorate is certainly becoming more polarized; according to Pew Research Center, the average partisan gap between voters, across 10 policy measures, is 36 percentage points as of 2017, up from 15 percentage points in 1994.^{8,9} Cohen, McGrath, Aronow, and Zaller 2016

5. Justin McCarthy, *Trump Seen as Less Conservative Than Prior GOP Candidates*, accessed December 7, 2019, <https://news.gallup.com/poll/196064/trump-seen-less-conservative-prior-gop-candidates.aspx>.

6. Randall G. Holcombe, *Public sector economics : the role of government in the American economy* (Pearson Education, 2006), 517, ISBN: 9780131450424.

7. Larry M. Bartels, *Opinion | The Rise of Presidential Extremists - The New York Times*, 2016, accessed December 7, 2019, <https://www.nytimes.com/2016/09/12/opinion/campaign-stops/the-rise-of-presidential-extremists.html>.

8. The ten policy positions poised to survey takers are: Government regulation of business usually does more harm than good, government is almost always wasteful and inefficient, poor people have it easy because they can get government benefits without doing anything in return, government today can't afford to do much more to help the needy, most corporations make a fair and reasonable amount of profit, blacks who can't get ahead in this country are mostly responsible for their own condition, immigrants today are a burden on our country because they take our jobs, housing and health care, homosexuality should be discouraged by society, the best way to ensure peace is through military strength, and stricter environmental laws and regulations cost too many jobs and hurt the economy.

9. *The Partisan Divide on Political Values Grows Even Wider | Pew Research Center*, accessed December 8, 2019, <https://www.people-press.org/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/>.

take this assessment and go further to attempt to explain this realized divergence from traditional spatial electoral theory, finding "scant evidence of extremism penalties that were either substantively large or close to statistical significance" for candidates or parties.¹⁰

One electoral feature at federal, state, and local levels in the United States, especially when considering polarization, is the primary election.¹¹ One possible thesis regarding political polarization places the blame on the primary system, an idea that has broad appeal among politicians. In a 2010 editorial in *The New York Times*, former Oregon Secretary of State Phil Keisling advocated "abolishing the primary system."¹² Current Senate Minority Leader Chuck Schumer (D-NY) wrote an op-ed for the same publication saying that, "the partisan primary system, which favors more ideologically pure candidates, has contributed to the election of more extreme officeholders and increased political polarization."¹³ Schumer is specifically rallying against the concept of a partisan primary, something that current Democratic contender for the nomination Bernie Sanders says posed a massive barrier to his candidacy due to his strong support among independents.¹⁴

There are numerous theoretical arguments to support the aforementioned beliefs.

For instance, those who vote in primaries could simply be more extreme as hypothesized by

10. Marty Cohen et al., "Ideologically Extreme Candidates in U.S. Presidential Elections, 1948–2012," ed. Larry M. Bartels, *The ANNALS of the American Academy of Political and Social Science* 667, no. 1 (September 2016): 126–142, ISSN: 0002-7162, doi:10.1177/0002716216660571, <http://journals.sagepub.com/doi/10.1177/0002716216660571>.

11. Ann OM. Bowman and Richard C. Kearney, *State and local government: the essentials* (Cengage Learning, 2015).

12. Phil Keisling, *To Reduce Partisanship, Get Rid of Partisans* - *The New York Times*, 2010, accessed December 8, 2019, <https://www.nytimes.com/2010/03/22/opinion/22keisling.html>.

13. Charles Schumer, *Opinion | Charles Schumer: Adopt the Open Primary* - *The New York Times*, 2014, accessed December 8, 2019, <https://www.nytimes.com/2014/07/22/opinion/charles-schumer-adopt-the-open-primary.html>.

14. Jeff Stein, *Bernie Sanders says Democrats should get rid of closed primaries. Is he right?* - *Vox*, 2016, accessed December 8, 2019, <https://www.vox.com/2016/4/28/11469468/open-primaries-closed-primaries-sanders>.

Jacobsen (2015); "Primary electorates are much more partisan and prone to ideological extremity."¹⁵ Sinclair (2006, 29-30) claims that activists "always vote, they vote in primaries, and they give money or work for candidates. To win in the primary, the candidate needs to be especially attentive and responsive to activists who are more polarized than less interested voters."¹⁶ Brady et al. examined primary and general election outcomes for Congress between 1956 and 1998 and found that "primaries pull candidates away from median district preferences."¹⁷ Burden (2004) found that candidates "who experience stiff primary competition are [...] more likely to deviate from the median voter's position."¹⁸

However, dominant evidence-based scholarship from more recent years largely doesn't support the view that primaries cause polarization. Hirano et al. 2010 examined the relationship between "the introduction of primary elections, the level of primary election turnout, or the threat of primary competition" and found limited evidence between those factors and polarization in roll call voting.¹⁹ McGhee et al. (2014) found that "the openness of a primary election has little, if any, effect on the extremism of the politicians it produces" by examining state legislator ideal points and matching them to the race's appropriate primary type.²⁰ Carson et al. (2018) derived measures of a congressperson's

15. Gary C. Jacobson, *The politics of congressional elections* (Rowman & Littlefield, 2015), 321, ISBN: 1442252634.

16. Barbara Sinclair, *Party wars: polarization and the politics of national policy making* (2006), doi:10.5860/choice.44-2391.

17. David W. Brady, Hahrie Han, and Jeremy C. Pope, "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?," *Legislative Studies Quarterly* 32, no. 1 (February 2007): 79–105, ISSN: 03629805, doi:10.3162/036298007X201994, <http://doi.wiley.com/10.3162/036298007X201994>.

18. Barry C. Burden, *Candidate Positioning in US Congressional Elections*, 2004, doi:10.1017/S000712340400002X.

19. Shigeo Hirano et al., "Primary elections and partisan polarization in the U.S. Congress," *Quarterly Journal of Political Science* 5, no. 2 (2010): 169–191, ISSN: 15540626, doi:10.1561/100.00008052.

20. Eric McGhee et al., "A Primary Cause of Partisanship? Nomination Systems and Legislator Ideology," *American Journal of Political Science*, 2014, ISSN: 15405907, doi:10.1111/ajps.12070.

ideology from campaign contributions and found that "unsuccessful challengers in the U.S. House are generally more extreme than those who win," and that our political institutions may be mitigating partisan polarization.²¹ Rogowski et al. (2015) used data from 85,000 major party candidates and found "no evidence that the restrictiveness of primary participation rules is systematically associated with candidate ideology."²² Woon (2018) conducted a controlled experiment which found that "primaries have little effect on average positions" of candidates.²³

This incongruity in the literature leaves open the question of the effect on partisan primaries on candidate positioning. While there is strong evidence that both support and argue against the view that primaries cause polarization, little research has been done on the effects of primaries on spatial positioning of candidates' communication. Existing literature also reveals a gap in the methodology used in the seminal literature on the topic. A possible avenue that has not yet been investigated when determining a candidate's positioning are their statements as made on Twitter, which has now become a dominant part of American political and electoral life.²⁴ Although there is some doubt regarding the veracity of using Twitter as an accurate metric of a candidates positioning, I will justify the use of Twitter in later chapters of the paper. I argue for a more limited scope than that advocated by the literature; candidates will moderate their expressed political lean on

21. Jamie L. Carson and Ryan D. Williamson, "Candidate ideology and electoral success in congressional elections," *Public Choice* 176, nos. 1-2 (July 2018): 175–192, ISSN: 15737101, doi:10.1007/s11127-017-0492-2.

22. Jon C. Rogowski and Stephanie Langella, "Primary Systems and Candidate Ideology: Evidence From Federal and State Legislative Elections," *American Politics Research*, 2015, ISSN: 15523373, doi:10.1177/1532673X14555177.

23. Jonathan Woon, "Primaries and Candidate Polarization: Behavioral Theory and Experimental Evidence," *American Political Science Review*, 2018, ISSN: 15375943, doi:10.1017/S0003055418000515.

24. Mike Isaac and Sydney Ember, "For Election Day Influence, Twitter Ruled Social Media," <https://www.nytimes.com/2016/11/09/technology/for-election-day-chatter-twitter-ruled-social-media.html>, (accessed: 06.06.2019).

Twitter to varying degrees depending on the type of primary they are engaged in. I hypothesize that open primaries will cause candidates to start closer to politically neutral and diverge from their initial positioning less over the course of the election when compared to their counterparts in closed primaries. This paper is not concerned with what kind of candidate each type of primary will produce, but rather how primary type will influence candidate moderation. To accomplish this, I will pull tweets from candidates for Senate in the 2014, 2016, and 2018 election seasons.²⁵ I propose a machine learning approach to categorizing tweets based on Kousser and Oklobdzija.

As such, the structure of this thesis will be as follows. Chapter 1 will discuss the history of the primary and the context under which this investigation arises, the motivation behind this investigation, Twitter as a medium, dominant literature within this topic, and two narrowly-tailored hypotheses for investigation and the relevant theoretical framework behind them. Chapter 2 will dive into the structure of the empirical investigation and justify the empirical decisions made throughout the process. Chapter 3 will discuss the results from the empirical investigation and place them within the context of the dominant literature. Chapter 4 will address weaknesses within the framework of the experiment and proceed to discuss routes of further investigation that could be undertaken by researchers to buttress the claims made in this paper.

25. Special elections are not included due to their abridged timeline. Several occurred for Senate in 2018.

Chapter 1

Relevance

Primaries are an extremely important feature of the United States electoral process and, according to Hirano et al. (2019), do several things that often go unnoticed.

1. They decide the winner of elections. Consider that among the 442 general elections for governor or United States senate held in the "Solid South" between 1878 and 1960, Republicans can only count one win.²⁶ Also consider that there are currently over 400 counties in the United States where the same party has dominated the presidential election for the past 50 years.²⁷
2. Primaries provide a forum for voters to evaluate candidates who are running apart from political divisions as, generally, candidates in a certain primary are from that primary's party. That narrows the ideological differences of candidates within a party versus between parties.
3. Primaries can serve as a debate stage for intra-party conflicts, allowing for open debates to settle the ideological differences within a party. Ideally, when the voters choose a nominee for a party, those differences will be resolved by the voters.

Because of these factors, it is important to understand the effect of primaries on the expressed ideology of politicians. To do this, we must first understand the context from which primaries were developed.

26. Shigeo Hirano and James M. Snyder, Jr, *Primary Elections in the United States* (Cambridge University Press, September 2019), ISBN: 9781139946537, doi:10.1017/9781139946537, <https://www.cambridge.org/core/product/identifier/9781139946537/type/book>.

27. Ibid.

History, Context, and Motivation

Before primaries, there existed the caucus-convention system.^{28,29} As opposed to modern primaries, which resolve themselves with a direct vote, the caucus-convention system acted in two stages. In the first stage, all constituencies involved in the election would select representatives to send to the convention. These delegates would then move to nominate candidates for different elected positions within their collective purview. Unfortunately, and as some scholars would argue, this system was extremely problematic and vulnerable to corruption on the part of party elites from the beginning.³⁰ On that subject, Peter Roberts, a Pennsylvania historian had the following to say:

Conventions, where nominations are made, have become in our counties bargain counters where candidates openly buy votes [...] Sums from \$75 to \$200 were recently paid and the men receiving the money attempted not to conceal the fact. An ex-boss in politics in one of our counties in a recent public address protested against bribery, corruption and commercialism in politics and said that the electors were sold “as sheep in the shambles.” A man in the crowd said: “He ought to know; a few years ago he brought into the county \$10,000 to buy up delegates for the state machine.”³¹

28. Frederick William Dallinger, *Nominations for Elective Office in the United States*, vol. 4 (Longmans, Green, / Company, 1897).

29. John F. Reynolds, *The Demise of the American Convention System, 1880–1911* (2009), doi:10.1017/cbo9780511511684.001.

30. Ralph S. Boots, “American Parties and Elections. BY Edward M. Sait. (New York: The Century Company. 1927. Pp. 608.),” *American Political Science Review* 21, no. 4 (September 1927): 900–902, ISSN: 0003-0554, doi:10.2307/1947614, https://www.cambridge.org/core/product/identifier/S0003055400024989/type/journal%7B%5C_%7Darticle.

31. C. R. Henderson, “Anthracite Coal Communities. Peter Roberts,” *American Journal of Sociology* 9, no. 6 (May 1904): 844–845, ISSN: 0002-9602, doi:10.1086/211277, <https://www.journals.uchicago.edu/doi/10.1086/211277>.

The corruption that was described by Roberts was not limited to Pennsylvania, but crossed state boundaries wherever there existed a caucus-convention system.³² The direct primary, in contrast, was seen as the antithesis of the caucus-convention system by putting the power of the nomination process with the people. This movement can primarily be credited to progressive leaders and part of a larger lobbying effort to pass a litany of reforms to weaken the power of political parties, bosses, and machines.³³

The goal of these primaries was to make elections more competitive by allowing intra-party challengers to enter the field. Indeed, seminal literature supports that this goal was realized for much of the 20th century. V. O. Key found there to be intense competition within the Democratic Primaries in the South at the beginning of the century, an example of advantaged-parties having more competitive races as a result of primaries.³⁴ Key also found a relationship between the prospects for victory and "the incidence of primary competition" in primaries for governor races between 1906 and 1952.³⁵ Hirano et al. (2019) suggests that this is true, finding that open-seat primaries for statewide offices were more likely to see increased competition from an advantaged party, with 84.5% of them being contested. The same pattern can be observed with House elections.³⁶ They also show that primary elections for statewide office were more important for deciding a winner than the

32. A.R. Conkling, "Reviews : Nominating Systems: Direct Primaries versus Conventions in the United States. By ERNST C. MEYER. Pp. 50l. Price, \$1.50. Published by the Author. Madison, Wisconsin, 1902," *The ANNALS of the American Academy of Political and Social Science* 21, no. 1 (January 1903): 112–114, ISSN: 0002-7162, doi:10.1177/000271620302100111, <http://journals.sagepub.com/doi/10.1177/000271620302100111>.

33. Gerald M. Pomper and Austin Ranney, "Curing the Mischiefs of Faction: Party Reform in America.," *Political Science Quarterly*, 1976, ISSN: 00323195, doi:10.2307/2149167.

34. Jay Topkis and V. O. Key, "Southern Politics in State and Nation," *The Yale Law Journal* 59, no. 6 (May 1950): 1203, ISSN: 00440094, doi:10.2307/793003, <https://www.jstor.org/stable/793003?origin=crossref>.

35. Wallace S. Sayre and V. O. Key, "American State Politics: An Introduction.," *Political Science Quarterly*, 1956, ISSN: 00323195, doi:10.2307/2145548.

36. Hirano and Snyder, Jr, *Primary Elections in the United States*.

general election when the incumbent was not running for re-election.

Much has changed from the early and mid 1900s, however, and one of those changes has come in the form of the electorate. According to the American National Election Studies, the percentage of independents from 1952 to 2008 has risen from 23% to 39%.³⁷ We can observe an incumbency advantage (due to candidate-centered versus party-centered voting) as a potential result of this change in the make-up of the population.³⁸ This shift has the potential to change the nature of primaries in modern elections. Further complicating this is the well-documented polarization of the parties.³⁹ As mentioned in the Introduction, primaries are a popular target for blame when it comes to polarization. This may be, in part, due to an ideological and physical "partisan sorting," whereby liberals go further to the left and conservatives go further to the right, entrenching them in the Democratic and Republican parties, respectively, while at the same time moving their geographic location, creating pockets of politically entrenched localities.⁴⁰ As a result of this, electoral environments have the potential to become highly polarized and primaries become more important for selecting the winner of an election than the general election.⁴¹

At the current date, primaries exist in several different forms: closed, semi-closed,

37. Marjorie Randon Hershey and John H. Aldrich, *Party politics in America* (2017), ISBN: 9781134836666, doi:10.4324/9781315544427.

38. Stephen Ansolabehere and James M. Snyder, "The Incumbency Advantage in U.S. Elections: An Analysis of State and Federal Offices, 1942–2000," *Election Law Journal: Rules, Politics, and Policy* 1, no. 3 (September 2002): 315–338, ISSN: 1533-1296, doi:10.1089/153312902760137578, <http://www.liebertpub.com/doi/10.1089/153312902760137578>.

39. Nolan McCarty, Keith T. Poole, and Howard Rosenthal, "The Polarization of the Politicians," in *Polarized America: The Dance of Ideology and Unequal Riches* (2006).

40. Wendy K. Tam Cho, *Voter migration is a significant factor in the geographic sorting of the American electorate* / *USAPP*, 2018, accessed December 9, 2019, <https://blogs.lse.ac.uk/usappblog/2013/11/19/voter-migration-is-a-significant-factor-in-the-geographic-sorting-of-the-american-electorate/>.

41. Scott Ashworth and Ethan Bueno De Mesquita, "Electoral selection, strategic challenger entry, and the incumbency advantage," *Journal of Politics*, 2008, ISSN: 00223816, doi:10.1017/S0022381608081024.

open, semi-open, blanket, and non-partisan.⁴² Each of these types has different implications in light of the aforementioned findings, and other state-by-state nuances to each of the primary types may slightly augment the effect of primaries across state boundaries. Of particular interest in this investigation are closed and open primaries which, when considering the increasing proportion of independents, could augment the ability of nationally moderate candidates to successfully win a nomination. The motivation of this investigation is to expand on the existing literature of primaries and partisanship by examining a novel dataset created using Twitter, providing potential insight through an avenue that has not yet been followed.

Literature Review

Theoretical Framework of the Median Voter Theorem in Elections

Can the median voter explain polarization? This question has been dissected by academics for potentially the last half century, at least since the theorem was formalized by Anthony Downs' *An Economic Theory of Political Action in Democracy*.⁴³

The Median Voter Theorem requires that candidates and voter preferences can be represented on a two-dimensional Euclidean space subject to several conditions. Gruber (2011) outlines the seven assumptions required for the theorem to hold true.⁴⁴ Although the number of assumptions isn't constant in the literature, the spirit and essence of those

42. *State Primary Election Types*, 2018, accessed December 10, 2019, <http://www.ncsl.org/research/elections-and-campaigns/primary-types.aspx>.

43. Anthony Downs, "An Economic Theory of Political Action in a Democracy," *Journal of Political Economy*, 1957, ISSN: 0022-3808, doi:10.1086/257897.

44. Jonathan Gruber, *Public Finance and Public Policy (3rd)* (Worth Publishers, 2011), 860, ISBN: 9781429219495.

assumptions is. For the purpose of this paper and simplicity's sake, we will be using the model as outlined by Gruber.

1. There is single dimensional voting; one issue at a time.
2. Voter preferences are single-peaked to prevent group preference cycling. This enforcement of a "pattern" violates Arrowian fairness and therefore ensures a consistent solution.
3. Voters are choosing between two options. Without this, there can be no equilibrium state of the game.
4. Politicians only care about winning.
5. There is no selective voting.
6. There are no external incentives such as lobbying.
7. There is no information deficit between parties.

These assumptions ensure a Condorcet winner, if it exists, will be selected through an election. One can observe below the way candidates or parties would interact with each other, keeping this in mind.

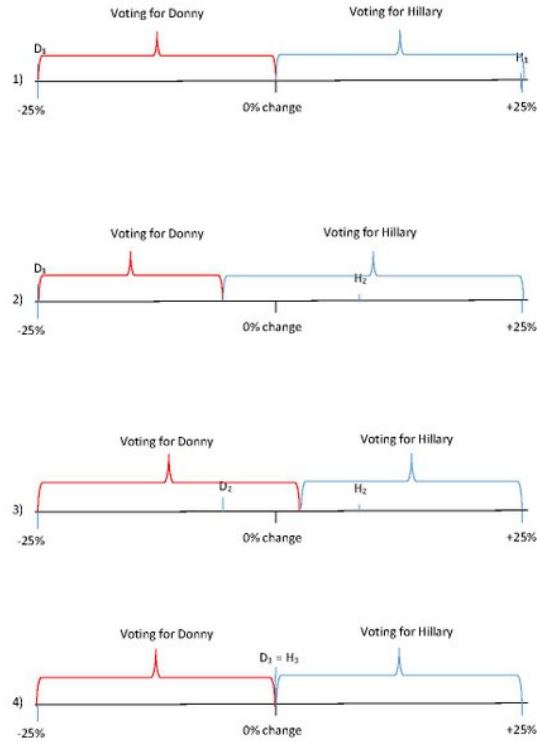


Figure 1: MVT example using the 2016 Presidential Election⁴⁵

Of special interest is the role of the median voter theorem in American politics as relates to the primary election process. Perhaps it is the introduction of this intermediate electoral mechanism that polarizes general elections. The same mechanic that is present in Figure 1 instead functions as displayed below when taking into account primary elections.

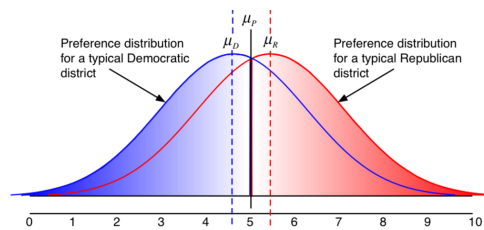


Figure 2: MVT with primary district-level preference overlaid. It is worth noting that party preferences would see distributions with medians closer to the periphery⁴⁶

It becomes clear, merely graphically, that the MVT could exacerbate partisanship as a result of primary elections. By selecting the median inter-group candidate for each party, partisan primary elections, so says the common wisdom, send candidates to the general who are far from the *national* median voter. The literature, however, is intensely divided.

The MVT can be proven within this theoretical framework. Supposing that there are two candidates, 1 and 2, and that they simultaneously announce and commit to policies p_1 and p_2 with the bliss point of the median voter b_{median} , we should observe that $p_1 = p_2 = b_{median}$. Suppose, then, that p_1 has a greater vote share than p_2 such that $p_1 < b_{median}$. This will force p_2 to moderate such that their position is an infinitesimally small ϵ greater than p_1 such that $p_2 < b_{median}$ changing voter preference from $[p_2, \infty)$ to p_2 . This leads to a cycle which only realizes equilibrium at $p_1 = p_2 = b_{median}$, proving the theorem.

There is some evidence of this theorem holding true in empirical studies. Holcombe's seminal paper on the topic in 1980 examined education expenditures for over 250 Michigan school districts and found that estimated levels of expenditures based on the Bowen equilibrium diverged by only 3% on average from actual expenditures.⁴⁷ Rice (1985) says that voters with a median income will vote for income redistribution via a tax hike on those above the median income, and shows a systematic closing of the gap between median and mean incomes.⁴⁸ This result helps to show MTV in action. Extending that, Husted and Kenny (1997) used MVT to explain the parallel rise in voting franchise and government

47. Holcombe, *Public sector economics : the role of government in the American economy*.

48. T. W. Rice, "An Examination of the Median Voter Hypothesis," *Political Research Quarterly* 38, no. 2 (June 1985): 211–223, ISSN: 1065-9129, doi:10.1177/106591298503800204, <http://prq.sagepub.com/cgi/doi/10.1177/106591298503800204>.

redistribution programs in the mid to late 1900s.⁴⁹ Fujiwara (2015) shows the theorem in action by looking at voting technology and political responsiveness in Brazil.⁵⁰ He shows that voting technology, which increased access to the ballot for the poor and less educated, resulted in increased government spending on health care and utilization of prenatal and newborn health services among less-educated women.

McKelvey and Ordeshook conducted a series of experiments that also lend support to the MVT and the potential for its observance in the real world even, when the pristine assumptions as outlined above are challenged. In 1982, they identified that a Condorcet winner will be converged on in two-dimensional policy spaces, much in the same way that a competitive market converges on an equilibrium price.⁵¹ However, this experiment was conducted with complete information about both candidate platforms and preference of voters by voters and candidates, respectively. They go on to assert in later research that complete information is not required for convergence on the Condorcet winner. In 1984, they used fulfilled rational expectations to conclude that perfect information is not necessary to observe the median voter theorem and confirm it for 1-dimensional electoral models between two candidates.⁵² They asserted that even with historical data as the only information available on policy positions of candidates, there is a notion of "stationary rational expectations equilibrium" which sees all voters optimize such that the resultant

49. Thomas A. Husted and Lawrence W. Kenny, "The effect of the expansion of the voting franchise on the size of government," *Journal of Political Economy* 105, no. 1 (October 1997): 54–82, ISSN: 00223808, doi:10.1086/262065.

50. Thomas Fujiwara, "Voting Technology, Political Responsiveness, and Infant Health: Evidence From Brazil," *Econometrica* 83, no. 2 (March 2015): 423–464, ISSN: 0012-9682, doi:10.3982/ecta11520.

51. Richard D. McKelvey and Peter C. Ordeshook, "Two-Candidate Elections without Majority Rule Equilibria," *Simulation & Games* 13, no. 3 (September 1982): 311–335, ISSN: 0037-5500, doi:10.1177/0037550082133004, <http://journals.sagepub.com/doi/10.1177/0037550082133004>.

52. Richard D. McKelvey and Peter C. Ordeshook, "Sequential Elections with Limited Information," *American Journal of Political Science* 29, no. 3 (August 1985): 480, ISSN: 00925853, doi:10.2307/2111140.

equilibria are the same as that which occurs with complete information. They also show this true for multiple issues and generalized preferences.⁵³ A 1985 paper from the same authors proved the median voter theorem robust for repeated elections and pure-strategy equilibria in both one and two dimensions with incomplete information.⁵⁴ A later paper used polling and interest groups as information transfer mechanisms and found that, using the rational expectations theory of markets, the aforementioned sources of information revealed information sufficient for voters to behave in a rational manner *as though they had complete information*.⁵⁵ The logical conclusion of that result is that candidates will, once again, converge on the median voter over time, and it can be shown to hold true in multiple dimensions albeit at a slower rate.

While McKelvey and Ordeshook above focused on candidate competition and under what conditions it will be observed, there is an equally important aspect of elections that they did not address; namely, will convergence on the median voter be realized based on the mechanisms and heuristics that voters use to vote?⁵⁶ More specifically, does retrospective voting result in convergence on the median voter's preference? It is more or less accepted by the literature that Americans vote retrospectively in contrast to European nations.⁵⁷ Campbell et. al. (2011) found that the theory of conditional voting holds true for

53. Richard D. McKelvey and Peter C. Ordeshook, "Elections with limited information: A multidimensional model," *Mathematical Social Sciences* 14, no. 1 (August 1987): 77–99, ISSN: 01654896, doi:10.1016/0165-4896(87)90016-3.

54. Richard D. McKelvey and Peter C. Ordeshook, "Rational expectations in elections: some experimental results based on a multidimensional model," *Public Choice* 44, no. 1 (January 1984): 61–102, ISSN: 00485829, doi:10.1007/BF00124819.

55. Richard D. McKelvey and Peter C. Ordeshook, "Elections with limited information: A fulfilled expectations model using contemporaneous poll and endorsement data as information sources," *Journal of Economic Theory* 36, no. 1 (June 1985): 55–85, ISSN: 10957235, doi:10.1016/0022-0531(85)90079-1.

56. McKelvey and Ordeshook effectively make voters use prospective voting as a result of the information sources they were provided i.e. polls, interest groups.

57. J. Edwin Benton, *When Americans go to the polls, they look to the past – not the future*, 2019, accessed February 25, 2020, <https://theconversation.com/when-americans-go-to-the-polls-they-look-to->

incumbents.⁵⁸ Lanoue (1994) finds that retrospective economic judgements are consistently powerful predictors of for voting behavior.⁵⁹ Hopkins and Pettingill (2018) found that retrospective voting also holds true for mayoral elections in large cities in the United States.⁶⁰

In answer to the question poised above, there is some evidence to indicate that median outcomes can be realized through retrospective voting alone. Collier et. al. (1987) addresses this question by conducting an experiment in which a sequence of elections is simulated between two candidates. Utilizing a one-dimensional policy space and single-peaked preferences, and without voter utility functions being known by the subjects, candidates are nevertheless found to converge on the median voter with voters only being aware of their payoff under the incumbent and the previous occupant of the position.⁶¹ What if voters are given the choice to collect more information at a cost and become better informed? Collier et al. (1989) addresses this question and finds that "voters purchase less information and rely more on retrospective knowledge."⁶²

From the preponderance of the literature, it would appear that we can hold as true two assumptions: the Condorcet winner predicts election outcomes well, and the median

the-past-not-the-future-114402.

58. James E. Campbell, Bryan J. Dettrey, and Hongxing Yin, "The theory of conditional retrospective voting: Does the presidential record matter less in open-seat elections?," *Journal of Politics* 72, no. 4 (October 2010): 1083–1095, ISSN: 00223816, doi:10.1017/S002238161000054X.

59. David J. Lanoue, "Retrospective and Prospective Voting in Presidential-Year Elections," *Political Research Quarterly* 47, no. 1 (March 1994): 193, ISSN: 10659129, doi:10.2307/448908.

60. Daniel J. Hopkins and Lindsay M. Pettingill, "Retrospective Voting in Big-City US Mayoral Elections," *Political Science Research and Methods* 6, no. 4 (October 2018): 697–714, ISSN: 20498489, doi:10.1017/psrm.2016.54.

61. Kenneth E. Collier et al., "Retrospective voting: An experimental study," *Public Choice* 53, no. 2 (January 1987): 101–130, ISSN: 00485829, doi:10.1007/BF00125844.

62. Kenneth Collier, Peter C. Ordeshook, and Kenneth Williams, "The rationally uninformed electorate: Some experimental evidence," *Public Choice* 60, no. 1 (January 1989): 3–29, ISSN: 00485829, doi:10.1007/BF00124309.

voter theorem is robust to changes in voter and candidate information possession. This is essential in order to extend these experimental proofs to real elections.

The Median Voter Theorem and Primaries, Reworked

There is a significant and shockingly intuitive argument against the above literature that relies on one of the fundamental assumptions of game theory: players act rationally. Serra (2015) and (2018) make this point regarding parties in primaries, arguing that if parties are actually completely rational actors, they will act prospectively and choose candidates best suited to win the general election.

Serra (2015) attempts to provide a definite answer to the conflicting results of empirical investigations as to whether primaries have statistically significant effects on polarization through the median voter theorem.⁶³ His model reveals something more extreme than prior papers; that there is not only no significant increase in polarization, but no increase in polarization at all when primaries are added to the traditional Downsian model. He models the election as a three-stage game with three players: voters, parties, and candidates. The game proceeds with an announcement of platforms by candidates, a nomination stage, and then a general election stage that decides the winning platform. Serra tests three so-called "centrifugal forces" through the assumptions of his model: parties have ideal points on opposite sides of the median voter, parties are policy, rather than office motivated, and platform promises have to be kept throughout the election and into office.

Taking voter preferences to be single-peaked, we know that voters will always vote

63. Gilles Serra, "No Polarization in Spite of Primaries: A Median Voter Theorem with Competitive Nominations" (2015), 211–229, doi:10.1007/978-3-319-15551-7_11, http://link.springer.com/10.1007/978-3-319-15551-7%7B%5C_%7D11.

for the policy position closest to their bliss point and traditional assumptions of the median voter theorem take hold; the candidate closest to the median voter will win the election. Parties are also policy motivated and each possess a single bliss point on opposite sides of the median voter. However, if parties are strategic, they will choose a candidate that generates the most utility for them over time, which is logically the candidate closest to the median voter of the general election that also maximizes their utility function with regards to the candidate's policy position. Absent outside influences, the median voter theorem results in centrist platforms regardless of competitive nomination stages. However, he points out that it may not be primaries themselves but conditions that act through the framework of primaries that cause polarization. If this model holds true, it is possible that empirical investigations flagging polarization are true as well, but as a result of an unidentified factor.

Serra (2018) adds two more centrifugal forces to the model as a robustness check against other potential causes of polarization, namely candidates who receive additional, independent payoffs for winning the primary, and risk-seeking parties who would prefer to nominate an extreme candidate instead of winning with a moderate candidate.⁶⁴ In both these new situations, polarization is still not realized. Serra concludes that it is the rationality of party members that is the most significant of these centripetal forces that drives platforms and candidates to the center, meaning that strategic voters will vote for a centrist candidate regardless of how extreme they actually are. Therefore even if the electorate is polarized, it is not the case that a strategic electorate would therefore choose an extreme candidate. That "strategic property" of voters may be a greater flag of

64. *Routledge Handbook of Primary Elections* (2018), doi:10.4324/9781315544182.

polarization through the framework of the primary than primary elections themselves.

Why Use Twitter?

We must examine why politicians themselves use Twitter to understand its purpose in answering questions such as those posed by this paper. Understanding how and why Twitter is distinct from other forms of political communication, including other forms of social media, and why politicians choose to use the platform over others, is critical for not only justifying its use in this empirical study, but also essential for the interpretation of any results that may reveal themselves as a result of the analysis herein performed.

Trends in recent years have indicated that there has been a significant widening of the gap between politicians and voters. Of particular interest is the concentration of that decline among the youngest of voters in the United States.⁶⁵ One possible theory put forth by Blais and Rubenson (2013) is that generational value shifts explain turnout declines.⁶⁶ They explain that "young voters are less inclined to vote because their generation is less prone to construe voting as a moral duty and is more skeptical about politicians' responsiveness to their concerns." Related to this, Anderson and McLeod (2004) discusses that politicians have lost touch with their constituents.⁶⁷ Although the aforementioned study focuses on the communication deficit of the European Parliament, the same effects

65. Andre Blais, Elisabeth Gidengil, and Neil Nevitte, "Where does turnout decline come from?," *European Journal of Political Research* 43, no. 2 (March 2004): 221–236, ISSN: 0304-4130, doi:10.1111/j.1475-6765.2004.00152.x, <http://doi.wiley.com/10.1111/j.1475-6765.2004.00152.x>.

66. André Blais and Daniel Rubenson, "The Source of Turnout Decline," *Comparative Political Studies* 46, no. 1 (January 2013): 95–117, ISSN: 0010-4140, doi:10.1177/0010414012453032, <http://journals.sagepub.com/doi/10.1177/0010414012453032>.

67. PETER J. ANDERSON and AILEEN McLEOD, "The Great Non-Communicator? The Mass Communication Deficit of the European Parliament and its Press Directorate*," *JCMS: Journal of Common Market Studies* 42, no. 5 (December 2004): 897–917, ISSN: 0021-9886, doi:10.1111/j.0021-9886.2004.00534.x, <http://doi.wiley.com/10.1111/j.0021-9886.2004.00534.x>.

can be observed in the United States. A 2015 Gallup report indicated that 48% of Americans say that their own representative is out of touch.⁶⁸ Congress itself also believes the country to be more conservative than it actually is. A 2013 working paper from UC Berkeley and University of Michigan researchers found that, among other things, congresspeople underestimate their constituents' support for universal healthcare by more than 15 points on average.⁶⁹ In 2018, a study from Columbia University and University of California Santa Barbara researchers found the same effects; representatives in Congress, for almost every policy item, assumed their constituents were far more conservative than reality.⁷⁰ It's no wonder, then, that public trust in government has gone down steadily, across party, racial, and generational lines for decades.⁷¹ Indeed, Twitter users tend to be "younger, more educated and more likely to be Democrats than the general public" according to a 2019 Pew Research study, bolstering its usefulness to combat the communication deficit politicians face.⁷²

In light of this, politicians appear to be using Twitter strategically to appeal to voters. LaMarre and Suzuki-Lambrech (2013) found that congressional campaigns successfully used Twitter to build relationship with democratically engaged citizenry and

68. Andrew Dugan, *Majority of Americans See Congress as Out of Touch, Corrupt*, 2015, accessed February 8, 2020, <https://news.gallup.com/poll/185918/majority-americans-congress-touch-corrupt.aspx>.

69. David E Broockman and Christopher Skovron, *What Politicians Believe About Their Constituents: Asymmetric Misperceptions and Prospects for Constituency Control* *, technical report (2013), <http://polisci.berkeley.edu/people/graduatestudents/persondetail.php?person=381>. <http://sitemaker.umich.edu/cskovron/home.1>.

70. Alexander Hertel-Fernandez, Matto Mildemberger, and Leah C. Stokes, *Opinion | Congress Has No Clue What Americans Want - The New York Times*, 2018, accessed February 8, 2020, <https://www.nytimes.com/2018/10/31/opinion/congress-midterms-public-opinion.html>.

71. Pew Research Center, *Public Trust in Government: 1958-2019* | Pew Research Center, accessed February 8, 2020, <https://www.people-press.org/2019/04/11/public-trust-in-government-1958-2019/>.

72. Stefan Wojcik and Adam Hughes, "Sizing Up Twitter Users," <https://www.pewinternet.org/2019/04/24/sizing-up-twitter-users/>, (accessed: 06.06.2019).

mobilize support for campaigns and candidates during the 2012 campaign.⁷³ Evans et al. (2014) found that, during the same campaign cycle, "incumbents, Democrats, women, and those in competitive races tweet differently than challengers, Republicans, minor party candidates, men, and those in safe districts."⁷⁴ A 2019 study looking at Switzerland, the UK, Italy, France, Denmark, and the United States found that populist politicians tend to use social media, including Twitter, to a greater extent than public appearances such as talk shows, and that the language found on social media is generally more populist than that used for public appearances.⁷⁵ As for outreach, there is some evidence that that 112th Congress, considered to be the first "tweeting" Congress, moderated style oriented to specific demographics in order to appeal to Latinos on Twitter.⁷⁶

Politicians also appear to be using Twitter intelligently. According to Bode et. al. (2015), clusters within candidate networks during the 2010 midterm elections expressed strategic hashtag usage. In particular, the political Right employs "hashjacking" in a more nuanced manner, where they co-opt hashtags used by political rivals in their own tweets in order to "maximize the diffusion of their views, either for proselytizing or to confront opponents."⁷⁷ The same phenomenon was observed independently among German

73. Heather L. LaMarre and Yoshikazu Suzuki-Lambrech, "Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns'," *Public Relations Review* 39, no. 4 (November 2013): 360–368, ISSN: 03638111, doi:10.1016/j.pubrev.2013.07.009.

74. Heather K. Evans, Victoria Cordova, and Savannah Sipole, "Twitter style: An analysis of how house candidates used twitter in their 2012 campaigns," *PS - Political Science and Politics* 47, no. 2 (2014): 454–462, ISSN: 15375935, doi:10.1017/S1049096514000389.

75. Nicole Ernst et al., "Populists Prefer Social Media Over Talk Shows: An Analysis of Populist Messages and Stylistic Elements Across Six Countries," *Social Media + Society* 5, no. 1 (January 2019): 205630511882335, ISSN: 2056-3051, doi:10.1177/2056305118823358, <http://journals.sagepub.com/doi/10.1177/2056305118823358>.

76. Bryan T. Gervais and Walter C. Wilson, "New media for the new electorate? Congressional outreach to Latinos on Twitter," *Politics, Groups, and Identities* 7, no. 2 (April 2019): 305–323, ISSN: 2156-5503, doi:10.1080/21565503.2017.1358186, <https://www.tandfonline.com/doi/full/10.1080/21565503.2017.1358186>.

77. Leticia Bode et al., "Candidate Networks, Citizen Clusters, and Political Expression," *The ANNALS of the American Academy of Political and Social Science* 659, no. 1 (May 2015): 149–165, ISSN: 0002-7162,

politicians by Darius and Stephany (2019), where they found that the German far-right party *Alternative für Deutschland* to intentionally polarize discourse, underlining the importance of online strategies as a tool for the recent success of far-right parties.⁷⁸

Interestingly, and more relevant to the discussion within this paper, Hemphill and Shapiro (2019) found that among incumbents running for re-election in 2016, Republicans exhibited partisan signalling leading up to the election, versus Democratic candidates who moderated their messaging on Twitter.⁷⁹

Past research on *traditional* negative campaigning found it to depress voter turnout and the electorate. Min (2004) found that personal attacks against opposing candidates "significantly depress one's participatory intent."⁸⁰ Djupe and Peterson (2002) found that divisiveness decreases turnout in senatorial primaries and weakens the divisive party's chances in the general election, based on an analysis of senatorial primaries in 1998.⁸¹ An analysis of the 2015 General Election in England similarly indicated that "electoral preferences are weakened for parties engaging in negative campaigning," going onto say that positive second-choice candidates can win out by being more positive than those beating them.⁸² Kahn and Jenny (1999) observed slightly more nuanced effects:

doi:10.1177/0002716214563923, <http://journals.sagepub.com/doi/10.1177/0002716214563923>.

78. Philipp Darius and Fabian Stephany, "'Hashjacking' the Debate: Polarisation Strategies of Germany's Political Far-Right on Twitter," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11864 LNCS (Springer, November 2019), 298–308, ISBN: 9783030349707, doi:10.1007/978-3-030-34971-4_21.

79. Libby Hemphill and Matthew A. Shapiro, "Appealing to the base or to the moveable middle? Incumbents' partisan messaging before the 2016 U.S. congressional elections," *Journal of Information Technology & Politics* 16, no. 4 (October 2019): 325–341, ISSN: 1933-1681, doi:10.1080/19331681.2019.1651685, <https://www.tandfonline.com/doi/full/10.1080/19331681.2019.1651685>.

80. Young Min, "News Coverage of Negative Political Campaigns," *Harvard International Journal of Press/Politics* 9, no. 4 (October 2004): 95–111, ISSN: 1081-180X, doi:10.1177/1081180X04271861, <http://journals.sagepub.com/doi/10.1177/1081180X04271861>.

81. Paul A. Djupe and David A. M. Peterson, "The Impact of Negative Campaigning: Evidence from the 1998 Senatorial Primaries," *Political Research Quarterly* 55, no. 4 (December 2002): 845–860, ISSN: 1065-9129, doi:10.1177/106591290205500406, <http://journals.sagepub.com/doi/10.1177/106591290205500406>.

82. Annemarie S. Walter and Cees van der Eijk, "Unintended consequences of negative campaigning: Back-

unsubstantiated attacks depress voter turnout, but legitimate criticisms increase turnout.⁸³ The same was observed by Fridkin and Kenney (2004) looking at 97 contested senate races across 1988, 1990, and 1992.⁸⁴ Griffin (2012) finds an interesting relationship between negativity and the party of the person who expresses it; while voter turnout goes up when neither Democrat nor Republican expresses negativity, it also increases when the Republican candidate is perceived as going negative, although the author addresses possible conflating reasons for this.⁸⁵

Nevertheless, the effectiveness of negative campaigning on Twitter is supported by the literature. Ceron and d'Adda (2015) found that negative campaigns on Twitter were effective, measured by unsolicited voting intentions, with stronger impacts while the attacker was themselves under attack.⁸⁶ In examining the 2016 Presidential Race on Twitter, Ross and Caldwell (2020) find that Donald Trump's rhetoric on Twitter supported the "going negative" strategy when directed at Hillary Clinton.⁸⁷ They highlight the important of textual features of his tweets, including capital letters and language choice, in

lash and second-preference boost effects in a multi-party context," *British Journal of Politics and International Relations* 21, no. 3 (August 2019): 612–629, ISSN: 1467856X, doi:10.1177/1369148119842038.

83. Kim Fridkin Kahn and Patrick J. Kenney, "Do Negative Campaigns Mobilize or Suppress Turnout? Clarifying the Relationship between Negativity and Participation," *American Political Science Review* 93, no. 4 (December 1999): 877–889, ISSN: 0003-0554, doi:10.2307/2586118, https://www.cambridge.org/core/product/identifier/S000305540021914X/type/journal%7B%5C_%7Darticle.

84. Kim Leslie Fridkin and Patrick J. Kenney, "Do Negative Messages Work?," *American Politics Research* 32, no. 5 (September 2004): 570–605, ISSN: 1532-673X, doi:10.1177/1532673X03260834, <http://journals.sagepub.com/doi/10.1177/1532673X03260834>.

85. Hannah Griffin, "Keep it Clean? How Negative Campaigns Affect Voter Turnout," *Res Publica - Journal of Undergraduate Research* 17, no. 1 (July 2012), <https://digitalcommons.iwu.edu/respublica/vol17/iss1/6>.

86. Andrea Ceron and Giovanna D'Adda, "E-campaigning on Twitter: The effectiveness of distributive promises and negative campaign in the 2013 Italian election," *New Media & Society* 18, no. 9 (October 2016): 1935–1955, ISSN: 1461-4448, doi:10.1177/1461444815571915, <http://journals.sagepub.com/doi/10.1177/1461444815571915>.

87. Andrew S. Ross and David Caldwell, "'Going negative': An APPRAISAL analysis of the rhetoric of Donald Trump on Twitter," *Language and Communication* 70 (January 2020): 13–27, ISSN: 02715309, doi:10.1016/j.langcom.2019.09.003.

addition to the use of the Twitter platform itself, as reasons for why going negative on Twitter is particularly effective. Gross and Johnson (2016) found that campaign negativity increases on Twitter as a campaign season progresses.⁸⁸ Bekafigo and Pingley (2017) analyzed Twitter use of gubernatorial candidates in 2011 and found that candidates go negative, in particular, when tweeting about policy, distinct from what is found in traditional media.⁸⁹ Such widespread use of the negative campaign tactic would not be present were it not indicative of electoral success on some level. This situates Twitter as a unique campaigning medium that may offer advantages to politicians.

Twitter is not just used to appeal to voters, however. It can also be used to coordinate disinformation campaigns and influence news media. Political rumoring during the 2012 election was studied by Shin et. al. (2016), finding that "rumor spreaders formed strong partisan structures in which core groups of users selectively transmitted negative rumors about opposing candidates."⁹⁰ Chong (2018) supports this, showing that homophily holds true for political hashtags within a Twitter network.⁹¹ Studying how politicians interact with journalists, Shapiro and Hemphill (2016) found that for key policy issues in 2013, policy around the economy, immigration, health care, and marginalized groups was

88. Justin H. Gross and Kaylee T. Johnson, "Twitter Taunts and Tirades: Negative Campaigning in the Age of Trump," *PS: Political Science & Politics* 49, no. 04 (October 2016): 748–754, ISSN: 1049-0965, doi:10.1017/S1049096516001700, http://www.journals.cambridge.org/abstract%7B%5C_%7DS1049096516001700.

89. Marija Bekafigo and Allison Clark Pingley, "Do Campaigns "Go Negative" on Twitter?," in *Politics, Protest, and Empowerment in Digital Spaces* (IGI Global, December 2017), 178–191, ISBN: 9781522518631, doi:10.4018/978-1-5225-1862-4.ch011, <http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-1862-4.ch011>.

90. Jieun Shin et al., "Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction," *New Media & Society* 19, no. 8 (August 2017): 1214–1235, ISSN: 1461-4448, doi:10.1177/1461444816634054, <http://journals.sagepub.com/doi/10.1177/1461444816634054>.

91. Miyong Chong, "Analyzing political information network of the U.S. Partisan public on twitter," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10766 LNCS (Springer Verlag, 2018), 453–463, ISBN: 9783319781044, doi:10.1007/978-3-319-78105-1_50.

congruent between Twitter posts from politicians and New York Times articles, although that relationship was negative when Democrats and Republicans were debating on that issue.⁹² When used for disinformation, these phenomenons and techniques are called "political astroturfing," defined as "a centrally coordinated disinformation campaign in which participants pretend to be ordinary citizens acting independently."⁹³

The ways in which Twitter augments campaigning as highlighted by the literature are invaluable to politicians and serve as some of the reasons that campaigns continue to take advantage of the platform.⁹⁴ It, in turn, serves as an invaluable tool for researchers to evaluate their political positioning and the positioning of their campaigns. Information communicated over Twitter is, for instance, perceived as equally persuasive and believable by the electorate as that which is communicated through traditional media platforms.⁹⁵ In particular, the generation of free media through Twitter due to extensive coverage of high-profile political tweets almost ensures that even non-users will consume Twitter-based election material,⁹⁶ something that the media⁹⁷ and academia⁹⁸ are aware of.

92. Matthew A. Shapiro and Libby Hemphill, "Politicians and the Policy Agenda: Does Use of Twitter by the U.S. Congress Direct <i>New York Times</i> Content?," *Policy & Internet* 9, no. 1 (March 2017): 109–132, ISSN: 19442866, doi:10.1002/poi3.120, <http://doi.wiley.com/10.1002/poi3.120>.

93. Franziska B. Keller et al., "Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign," *Political Communication*, October 2019, 1–25, ISSN: 1058-4609, doi:10.1080/10584609.2019.1661888, <https://www.tandfonline.com/doi/full/10.1080/10584609.2019.1661888>.

94. Michael Bossetta, "The Digital Architectures of Social Media: Comparing Political Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election," *Journalism & Mass Communication Quarterly* 95, no. 2 (June 2018): 471–496, ISSN: 1077-6990, doi:10.1177/1077699018763307, <http://journals.sagepub.com/doi/10.1177/1077699018763307>.

95. David S. Morris, "Twitter Versus the Traditional Media," *Social Science Computer Review* 36, no. 4 (August 2018): 456–468, ISSN: 0894-4393, doi:10.1177/0894439317721441, <http://journals.sagepub.com/doi/10.1177/0894439317721441>.

96. Peter L. Francia, "Free Media and Twitter in the 2016 Presidential Election," *Social Science Computer Review* 36, no. 4 (August 2018): 440–455, ISSN: 0894-4393, doi:10.1177/0894439317730302, <http://journals.sagepub.com/doi/10.1177/0894439317730302>.

97. Akintunde Ahmad et al., *Sleepwalking into 2020 - Columbia Journalism Review*, 2020, accessed February 22, 2020, https://www.cjr.org/special%7B%5C_%7Dreport/2020-election-media-journalism.php.

98. John Bryden and Eric Silverman, "Underlying socio-political processes behind the 2016 US election,"

Hypotheses

After an analysis of the significant literature surrounding primary elections and polarization, I have selected the below hypotheses for investigation on Twitter.

1. Republicans and Democrats will each tweet to their respective bases more during the primary versus during the general election.
2. The above effect will be magnified in closed primary races when compared to open primary races.

I justify the above hypotheses with the preponderance of the traditional literature regarding both primary elections and polarization and Twitter as a platform itself.

H1 is stepped in the traditional game-theoretical model of spatial elections that has been rigorously proven by a multitude of studies that verify its veracity and showcase that it is robust to changing conditions of an election. Holcombe (1980) is the seminal empirical study that showed near insignificant deviation from the Bowen equilibrium on average in education expenditures.⁹⁹ Rice (1985) later justified the theorem with real outcomes that resulted from an expanded electorate, showing median policy positions were observed due to a shift in the composition of the electorate,¹⁰⁰ and the results were bolstered by Husted and Kenny (1997)¹⁰¹ and Fujiwara (2015)¹⁰² in different political contexts. The United

States' primary elections supposedly see this shift in the composition of the electorate even

ed. Haroldo V. Ribeiro, *PLOS ONE* 14, no. 4 (April 2019): e0214854, ISSN: 1932-6203, doi:10.1371/journal.pone.0214854, <http://dx.plos.org/10.1371/journal.pone.0214854>.

99. RANDALL G. HOLCOMBE, "AN EMPIRICAL TEST OF THE MEDIAN VOTER MODEL," *Economic Inquiry* 18, no. 2 (April 1980): 260–274, ISSN: 14657295, doi:10.1111/j.1465-7295.1980.tb00574.x, <http://doi.wiley.com/10.1111/j.1465-7295.1980.tb00574.x>.

100. Rice, "An Examination of the Median Voter Hypothesis."

101. Husted and Kenny, "The effect of the expansion of the voting franchise on the size of government."

102. Fujiwara, "Voting Technology, Political Responsiveness, and Infant Health: Evidence From Brazil."

beyond mere party affiliation; Jacobsen (2015) asserts that primary elections contain more partisan and extreme voters,¹⁰³ and Sinclair (2006) claims that activists form a larger portion of the voting electorate in primaries.¹⁰⁴ This should indicate a strong observable effect.

H1 is only possible because the median voter theorem has been shown to be robust to change by McKelvey and Ordeshook in their series of studies throughout the 1980s.¹⁰⁵ Their studies serve as landmark indications that even with interference in the form of incomplete information, multiple issues, generalized preferences, repeated elections, and interest groups as information transfer mechanisms, median outcomes are the only realized scenarios as people act as though they had complete information anyway.

Why would this be observed on Twitter specifically? LaMarre (2013) indicated that congressional campaigns used Twitter to build relationships with the electorate during the 2012 campaign season.¹⁰⁶ It stands to reason that because of this, communication between politicians and Twitter users would accurately reflect their platforms. There is also evidence that candidates use Twitter strategically depending on their constituencies and personal identities.¹⁰⁷ Twitter itself has seen a growth of types of political communication, including negative campaigning and disinformation campaigns, rendering it a sophisticated interface between candidates and voters.

With the results of the above literature, it seems logical that candidates will use

103. Jacobson, *The politics of congressional elections*.

104. Sinclair, *Party wars: polarization and the politics of national policy making*.

105. See the Literature Review for a complete summary and citations.

106. LaMarre and Suzuki-Lambrecht, "Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns'."

107. Evans, Cordova, and Sipole, "Twitter style: An analysis of how house candidates used twitter in their 2012 campaigns."

Twitter as a mechanism to communicate with their respective bases throughout the election. Applying the spatial elections model to this assumption, it would stand to reason that the effects would be stronger during the primary election where the electorate is more or less restricted to the candidates' own party. That trend, then, should reverse during the general. H2 then follows logically with the assumption that politicians know the rules of the election that they are in. With the knowledge that a primary is open, we should see candidates moderate more and express less extreme positions as party-members are not the only voters who they may be seeking to court early on. We would expect to see the opposite for closed primaries.

How do these hypotheses respond to the dominant recent scholarship that contends these effects do not exist? Hirano et al. (2010) finds little evidence that primaries are electing polarized officials,¹⁰⁸ and McGhee et al. (2014) challenges H2, finding that the degree of openness of a primary doesn't have a detectable effect on the extremism of the elected politician.¹⁰⁹ Carson et al. used campaign contributions to map candidates ideologically and saw that it was the extreme candidates, as rated by their campaign contributors, who lose.¹¹⁰ Rogowski et al. (2015) found the same through a larger scale analysis and saw no change in candidate ideology (as measured by Bonica's (2014) Database on Ideology, Money in Politics and Elections (DIME)) and primary restrictiveness.¹¹¹

However, these studies all do one of two things: either they measure the ideology of

108. Hirano et al., "Primary elections and partisan polarization in the U.S. Congress."

109. McGhee et al., "A Primary Cause of Partisanship? Nomination Systems and Legislator Ideology."

110. Carson and Williamson, "Candidate ideology and electoral success in congressional elections."

111. Rogowski and Langella, "Primary Systems and Candidate Ideology: Evidence From Federal and State Legislative Elections."

winning elected officials and peg them against the rules of the primary they went through, or they measure ideology through campaign contributions or roll calls. The first approach is distinctly different from what is to be done in this paper; neither hypothesis asserts that the result of the election will be a more polarized Senate, but rather that the primary elections, moreso than the general elections, will see more polarized communication across Twitter, and that those observable effects will be stronger in closed primary states when compared to open primary states. I make no claim regarding the actual positioning of these candidates after being elected, but rather only discuss their spatial positioning as seen through their communication on Twitter throughout the campaign. The second approach is also different because it aims to evaluate the candidates through a lens detached from political communication. Campaign contributions and roll call votes are two methods of ideological evaluation that are wonderful at placing politicians on an ideological mapping that is attached to their "true" positioning, but have nothing to do with how they communicate. Therefore I don't see this literature as challenging the veracity of my hypotheses.

What then of Serra (2015) and (2018) which attack the theoretical basis of my hypotheses? His argument is that candidates do not have to change their positioning based on whether it is the primary or general election because an electorate that is truly rational would not pick someone unlikely to win the general election (i.e. someone who would need to change their positioning). While I agree that this theory would likely predict the winner of a given election, I don't think that it significantly changes the thesis as relates to political communication because of what we observe with the dominant evidence-based literature in political communication: for instance, Acree et al. (2020) observes the exact predicted

sinusoidal movement of candidate positioning from the primary to the general when looking at Presidential elections.¹¹² Burden (2004) sees the same for Congressional elections.¹¹³

While I acknowledge the theoretical challenge, I don't see it to be stronger than the empirical results that support the traditional theoretical framework, and suspect that the evidence will vindicate that. It is important to note here, however, that Serra does argue that his theoretical framework and the literature I cite here can simultaneously be true as a result of outside contributing factors that cause polarization within primaries rather than primaries themselves being the cause. I leave that investigation to future studies.

112. Brice D.L. Acree et al., "Etch-a-Sketching: Evaluating the Post-Primary Rhetorical Moderation Hypothesis," *American Politics Research* 48, no. 1 (January 2020): 99–131, ISSN: 1532-673X, doi:10.1177/1532673X18800017, <http://journals.sagepub.com/doi/10.1177/1532673X18800017>.

113. Burden, *Candidate Positioning in US Congressional Elections*.

Chapter 2

Methods

The method of investigation for this study is supervised machine learning, which, in this case, takes advantage of the advanced predictive ability of computer algorithms and harnesses them for a text classification exercise. By taking inputs that are similar to target text, in this case tweets by politicians, and human-coded categorizations for that text, a classification algorithm can be trained to recognize traits that are associated with a particular category and predict the category of new text it has never seen before.

Supervised learning models are widely used in data science and consumer applications in every day life, including from benign situations such as e-mail spam filters to complex scenarios such as computer vision. Text classification specifically (of which e-mail spam filters are one application) allows algorithms to take text as input and sort them into any coded category the algorithm has been trained to recognize. Although to my knowledge it has never been applied to an investigation of these questions, the application of supervised learning to Twitter within political science is widespread in both the United States and abroad and has been widely published wherever the research takes place. Figure 3 below explains the process visually.

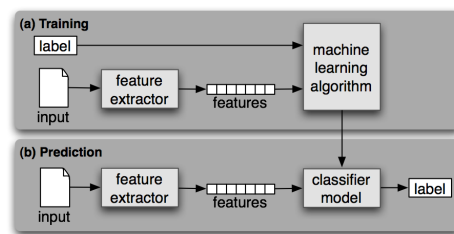


Figure 3: Diagram of Supervised Machine Learning

A significant disadvantage of a supervised learning approach is that a sufficiently large corpus of pre-categorized data is required to train a good algorithm. This corpus is called a "training set" and is fed to an algorithm in order to teach it how to sort input data. The training set used for the algorithms was created from the tweets of 23 presidential candidates' Twitter accounts during the 2016 Presidential Election provided by Professor Thad Kousser in the Department of Political Science at the university. It contains 7,525 human-coded tweets spanning from October 2015 to January 2017. Beginning in June 2016, six undergraduate and graduate research assistants categorized tweets on twelve separate measures, one of which is used here, receiving only the text of the tweet and a codebook that described what each specific coding might look like. Intercoder reliability metrics for Predicted Ideology (the variable of interest), calculated through an analysis of 1,217 tweets which were assigned to overlapping pairs of coders, are reported in Appendix 1. Agreement rate is the percentage of the time that human coders agreed with each other. Cohen's Kappa is a measure that attempts to account for the potential that some of the agreement rate is due to random chance. Higher is better for both of these metrics. Once the training set was created, I used it to train six types of algorithms, each taking one or more of four potential kinds of inputs, and evaluated the overall accuracy and Cohen's Kappa of each. A full listing of the measures and the algorithms used, their accuracy, and the corresponding Cohen's Kappa, can be found in Appendix 1.

In order to proceed with this investigation, I required a corpus of tweets from every candidate for the Democratic and Republican nomination for Senate in each state with a Senate election during 2014, 2016, and 2018. I sourced the names of these candidates from FEC filings and only selected those candidates that appeared on the ballot for the primary

Year	Start Date	End Date
2014	7 February 2014	4 November 2014
2016	6 November 2015	8 November 2016
2018	4 December 2017	6 November 2018

Table 1: Start and End Dates for Data Collection By Election Year

election. I then cross-referenced their names with accounts on Twitter and determined their Twitter presence, if any. If they had multiple accounts, I picked the one with a greater following. I then used the online tool Crimson Hexagon to export the tweets for the election period tracked. Table 1 contains the dates tracked for each election.

After downloading all of the candidate tweets to individual CSV files, I then cleaned the tweets using standard natural language processing (NLP) practices, including stemming, removing stop words (determiners, coordinating conjunctions, prepositions), numbers, special characters and punctuation (including hashtags, although not their core phrases) and clearing capitalization of letters. This text cleaning is an essential part of any study involving natural language processing as it removes traits that are too complex for a computer to understand. By cleaning text, we allow supervised learning algorithms to accurately attribute a particular categorization to specific traits within a document, in this case a specific tweet. For instance, #MakeAmericaGreatAgain is a clear conservative flag. By removing irrelevant text, such as stop words and specific numbers, we can isolate those key phrases and improve classification accuracy. At the same time, we can keep track of these attributes, such as the number of capital letters, total words/characters of the original tweet, etc., and control for them in a final regression. An example is presented in Figure 4.

After cleaning the text of the tweets, I fed them through eleven separately trained supervised machine learning algorithms which categorized tweets on eleven different

Original v. Clean Tweet (@TeamMitch, 2014 Election)
When it comes to coal, @billclinton and @AlisonForKY are the same as @Barack-Obama! #kysen http://t.co/MnvuHDOCZ8
come coal billclinton alisonforky barackobama

Figure 4: Example of Tweet Cleaning

measures, one measure each. The most important measure is the Predicted Ideology (*IdeologyPred*) measure, which represents the predicted ideology on a liberal-neutral-conservative scale, represented by 0-1-2 values respectively. This variable will be the variable of inquiry for my hypotheses. Once the tweets were appropriately sorted, I combined all the sub-sets from candidates into a massive dataset containing all tweets sent by candidates for U.S. Senate during the 2014, 2016 and 2018 election periods tracked. At this point, I identified users who were outliers by measure of average posts per week and excluded them from the dataset so they would not bias the results.¹¹⁴ I also removed accounts that presented data challenges.

Data Inspection

An interesting observation from the complete dataset is that the mean ideology score for both Democrats and Republicans would indicate they tweet conservatively on an objective scale. Assuming the algorithm is sorting them correctly, Republicans and Democrats are ideologically close in their communication on average through election periods. Table 4 goes into more detail. However, the variation between the parties across the three elections, on a week-by-week basis, conforms to expected trends (Figure 4). Figure 5 is a trend line of

¹¹⁴. Outliers were trivially obvious, with often over one hundred or two hundred thousand tweets over the course of a one-year election cycle.

that weekly difference in ideology between the two parties.

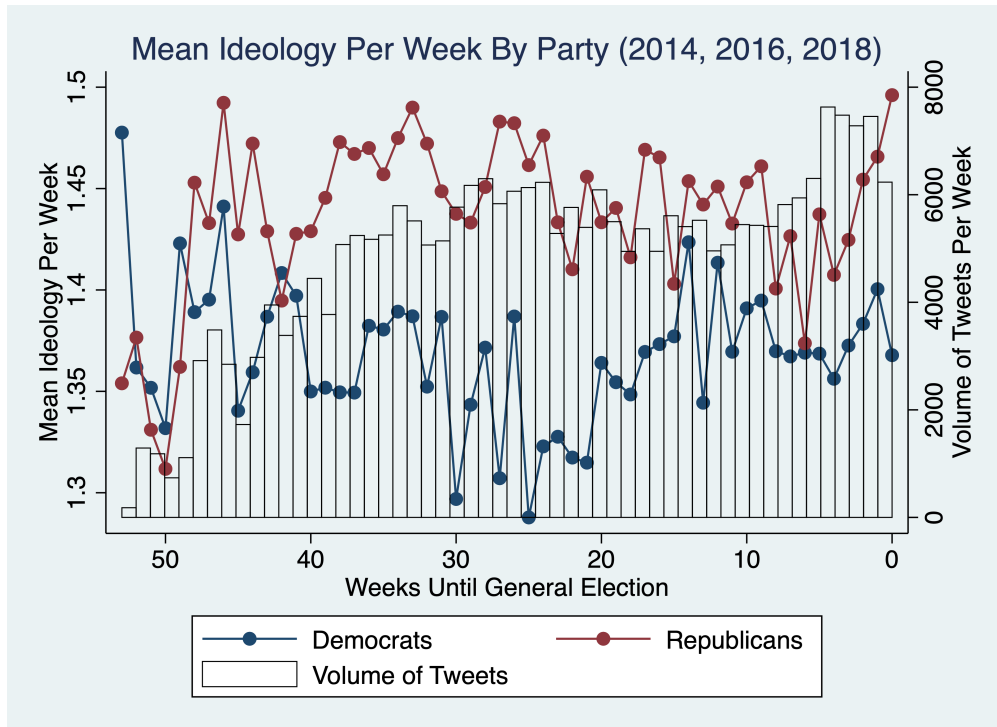


Figure 5

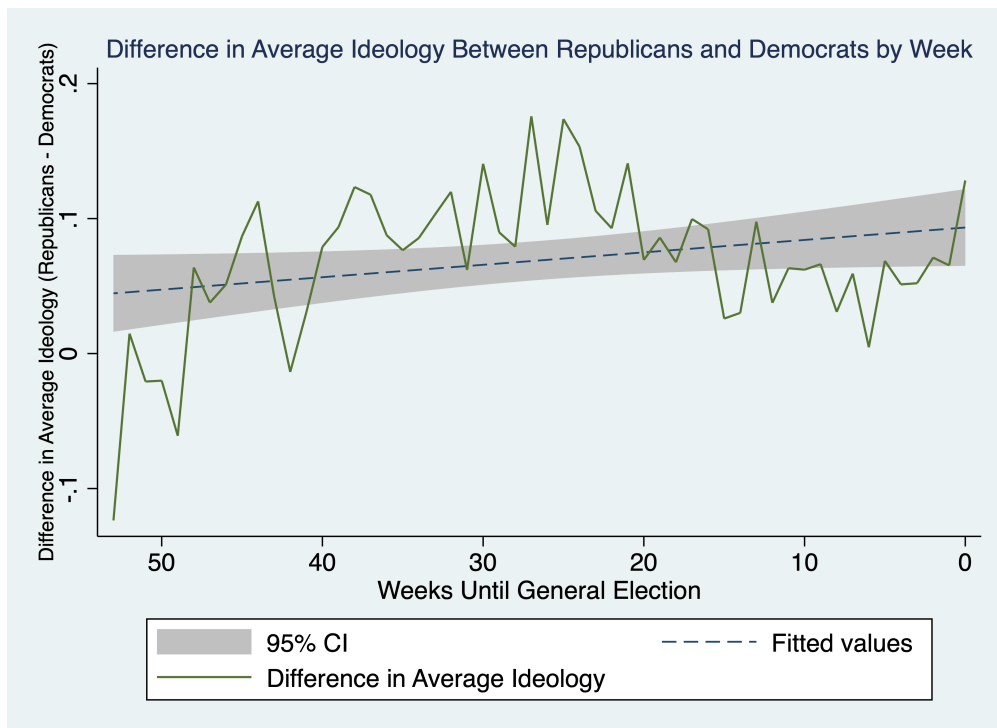


Figure 6

Year	Number of Tweets	Mean Predicted Ideology	Observed Candidates	% Total Candidates
2014	60,939	1.476	92	40.17%
2016	75,167	1.325	73	32.59%
2018	126,674	1.420	96	42.66%
Total	262,780	-	255	37.61%
Mean	87,593	1.405	85	38.47%
Democrats	127,363	1.365	101	37.41%
Republicans	135,417	1.445	155	37.99%

Table 2: Summary Statistics of the Final Data Set

Candidate_State Date	Cleaned Tweet Text	Predicted Ideology	Original Tweet
Begich_AK 02/11/2014	tough loss kikkanimal everywhere still proud hard olympics usa	1	https://bit.ly/3djlR94
Wicker_MS 06/11/2018	past week marked 500days realdonaldtrumps presidency weve delivered many promise im looking forward success come httpstcokjkd-fgmn2o wickerreport	2	https://bit.ly/394It0I
Sanders_VT 10/18/2018	youre literally working alongside 20 republican attorney general end protection preexisting condition one estimate many thousand die republican friend successful	0	https://bit.ly/33G38HY
Hawley_MO 10/14/2018	chucktodd confirmation process missourian cannot believe conduct democrat adding seeing mob behavior across country mtp	2	https://bit.ly/3diWxrh
Isakson_GA 05/06/2016	honor thank milspouses support troop militaryspouseappreciationday httpstco08jjuqlmpa	1	https://bit.ly/2xZ70wJ

Table 3: Sample Tweets

Party (Election)	2014	2016	2018	Means
Democrats (Primary)	1.379	1.303	1.363	1.355
Republicans (Primary)	1.542	1.336	1.487	1.442
Democrats (General)	1.443	1.329	1.381	1.381
Republicans (General)	1.556	1.322	1.497	1.496
Means	1.476	1.325	1.420	1.406

Table 4: Summary Statistics of Ideology by Party & Election

Of note is the divergence between the parties during the middle of the election season and the resultant merging of their ideologies approaching the General. The appearance of this phenomenon is a good indication that the model is functioning as expected regardless of the unexpected small differentiation in the average ideology scores between the two parties, and it is well-grounded in the literature. Polsby and Widavsky (1976)¹¹⁵ put forth the idea that candidates move ideologically in a sinusoidal fashion, going to the extremes during the primary and then shifting back to the center during the general election in order to maximize their vote share. Although their book discussed presidential elections, similar research has been done by Burden (2004) as discussed in the Literature Review.¹¹⁶ Acree et. al. (2020) supports the methodology of supervised machine learning to evaluate candidate positioning and agrees with the preliminary results here, finding evidence that presidential candidates in the 2008 and 2012 elections acted in a manner consistent with that described in Figure 5.¹¹⁷ Potential reasons for the small size of the deviation between ideology scores are put forth in later sections.

After the dataset of candidate tweets was completed, I then matched candidates to their states' primary rules using two methods of classification. The first is a granular

115. *Presidential Elections: Strategies and Structures of American Politics* (1976).

116. Burden, *Candidate Positioning in US Congressional Elections*.

117. Acree et al., "Etch-a-Sketching: Evaluating the Post-Primary Rhetorical Moderation Hypothesis."

classification of primary rules obtained from the National Conference of State Legislatures, which includes 5 distinct categories for state primary election types: closed, partially closed, partially open, open to unaffiliated voters, and open primaries.¹¹⁸ California and Washington use top-two primaries in which all candidates are listed on the same ballot. They have been coded as a separate kind of primary election. I did the same for Louisiana and Nebraska; Louisiana uses a top-two initial stage and a potential runoff election follows if no candidate gets more than 50% of the vote, and Nebraska holds non-partisan elections in which all candidates appear on the same ballot but without party designations attached to them. None of these four states will be included in the general analysis, but rather warrant separate analysis outside the scope of this paper. The second, less granular classification I used only differentiates between open and closed primaries. Closed and partially closed primaries are included in the latter, with partially open, open to unaffiliated voters and open primaries in the former. Table 5 contains a complete listing.

118. *State Primary Election Types*.

Primary Type	Primary Group	States
Closed	Closed	Delaware, Florida, Kentucky, Maryland, Nevada, New Mexico, New York, Oregon, Pennsylvania
Partially Closed	Closed	Alaska, Connecticut, Idaho, North Carolina, Oklahoma, South Dakota, Utah
Partially Open	Open	Illinois, Indiana, Iowa, Ohio, Tennessee, Wyoming
Open to Unaffiliated Voters	Open	Arizona, Colorado, Kansas, Maine, Massachusetts, New Hampshire, New Jersey, Rhode Island, West Virginia
Open	Open	Alabama, Arkansas, Georgia, Hawaii, Michigan, Minnesota, Mississippi, Missouri, Montana, North Dakota, South Carolina, Texas, Vermont, Virginia, Wisconsin
Top-Two	-	California, Washington
Other	-	Louisiana, Nebraska

Table 5: Primary Types and Their Respective States

Chapter 3

Results

This paper poses two hypotheses for investigation:

1. Republicans and Democrats will each tweet to their respective bases more during the primary versus during the general election
2. The above effect will be magnified in closed primary races when compared to open primary races

The primary variable of interest will be Predicted Ideology (IdeologyPred), which is a measure of ideology on a discrete one-dimensional spectrum from 0-1-2, representing liberal-neutral-conservative, respectively. Predicted Ideology over time can be used in order to evaluate the statistical significance of the hypotheses. In order to do this, I will be running bi-variate and multi-variate regressions, both with and without candidate-level fixed or random effects depending on what analysis of the data yields, for the tweets of both Democratic and Republican candidates as follows.

I will proceed with H1 by running both bi-variate and multi-variate regressions in order to isolate the effect of campaign timeline, as defined by a binary variable indicating whether the primary is happening right now or not, on Predicted Ideology of tweets sent by candidates within the two parties. I will run both regressions for each party alone and control for tweet attributes in the multi-variate regression. Based on the results of a Hausman specification test, I will then decide whether to use a fixed or random effects model for both bi-variate and multi-variate regressions for each party in order to counter

potential omitted variable bias. I will then analyze the totality of the results in light of the claim in my hypothesis and take a position on whether or not it has been robustly proven within this context.

I will proceed with H2 in a similar fashion with the same litany of regression tools. While both the dependent variable, independent variable, and controls will be the same, regressions will not only be split by party but also by primary type. This will allow us to compare the regression coefficient for the effect of a primary, by party, depending on the openness of that primary, in order to evaluate the claim made.

I use linear regression here because of its analytic simplicity and because inspection of the data, dominant literature, and intuition all suggest a broadly linear relationship between Predicted Ideology and election timeline. I find it unlikely that a linear model would not be able to fit the data reasonably well.

H1

For each of the candidates, the complete dataset contains, among other variables, a Predicted Ideology measure for each tweet and a binary variable that indicates whether the tweet was posted during the primary period of that candidate's state. Table 6 contains an abbreviated sample of some data including the main dimensions used in the regression. The candidates' party affiliations are also indicated.¹¹⁹

I first conducted a bi-variate regression with Predicted Ideology as the independent variable and a binary dependent variable indicating whether the primary was happening at

119. Senator Bernie Sanders is a declared independent but caucuses with the Democrats. Senator Angus King is a declared independent but caucuses with the Democrats. All candidates for the Democratic nomination in Minnesota appear as Democratic-Farmer-Labor Party candidates, but are Democrats for the purposes of this paper.

the time the tweet was posted or not. The results of this regression are presented in Table 7. Initially, they are concerning. Although Democrats appear to have the expected response to the stimulus of primaries, that is to say that they are more liberal during primary elections when compared to the general, Republicans appear to move their ideology slightly to the left during primaries as well, which is intuitively at odds with what we saw in Figure 4. However, tweet attributes have not been controlled for.

Table 8 reveal what are likely more accurate results.¹²⁰ Indeed, the effect of primaries is now in the right direction for both parties. One interesting thing to note is the success of the number of hashtags for predicting ideology among Republicans. This could be an indication that the model is picking up on hashjacking as described by Bode et. al. (2015) and Darius and Stephany (2019).¹²¹¹²² Both papers found that the political right uses hashtags, and the hashjacking technique in particular, far more than the left.

What about omitted variables? Certainly there are traits that could have been omitted from the dataset that are significant, such as gender, age, race competitiveness, population demographics of the voters, etc. It is possible that these variables would need to be controlled for, but as a consequence of time, the dataset does not contain data regarding any of the aforementioned measures. Based on this, a fixed or random effects model at the candidate-level may be appropriate.

In order to do this, I first collapsed the complete dataset so that tweet ideology was averaged per day per candidate in order to format the panel acceptably for regression. In

120. Hashtags and numerics were excluded from the Democrat regression as they have no predictive value for their tweets.

121. Bode et al., “Candidate Networks, Citizen Clusters, and Political Expression.”

122. Darius and Stephany, ““Hashjacking” the Debate: Polarisation Strategies of Germany’s Political Far-Right on Twitter.”

order to determine whether to use a fixed or random effects model, I then ran a Hausman specification test on the fixed and random effect models for both parties which confirmed that fixed effect models would be better used. This has some support in the literature from Burden (2004), which suggested that there are variations in candidate positioning depending on observable factors such as public reputation, weak general election competition, and stiff primary competition.¹²³ However, as Table 9 indicates, candidate fixed effects regressions do not result in statistically significant effects.

What about multi-variate candidate level fixed- or random-effects? The Hausman specification test once again supported a fixed effects model, the results of which can be found in Table 9 as well. Interestingly, Democratic fixed effects, controlling for relevant tweet attributes, does result in a weakly statistically significant result.

The mixed results allow us to conclude a few things. First, H1 is weakly proven across candidates by virtue of the multi-variate models for each party. Statistically significant effects in the directions that the hypothesis predicts are present for both Democrats and Republicans. However, candidate-level fixed-effect models reveal that while this trend also functions within individual Democratic candidates, the same is not true for Republican candidates. There are some potential explanations as to why this is. The most obvious is that the fixed effects model is not enough to account for the omitted variable bias for the Republican Party. The dataset may not contain enough data at a candidate level for intra-candidate effects to be observed. Another potential issue is data availability by candidate as many candidates did not tweet much at all over the course of the election.

123. Burden, *Candidate Positioning in US Congressional Elections*.

Candidate_State	Date	Cleaned Tweet Text	Predicted Ideology	Is the Primary Now?
Begich_AK	02/11/2014	tough loss kikkanimal everywhere still proud hard olympics usa	1	1
Wicker_MS	06/11/2018	past week marked 500days realdonaldtrumps presidency weve delivered many promise im looking forward success come httpstcokjkd-fgmn2o wickerreport	2	1
Sanders_VT	10/18/2018	youre literally working alongside 20 republican attorney general end protection preexisting condition one estimate many thousand die republican friend successful	0	0
Hawley_MO	10/14/2018	chucktodd confirmation process missourian cannot believe conduct democrat adding seeing mob behavior across country mtp	2	0
Isakson_GA	05/06/2016	honor thank milspouses support troop militaryspouseappreciationday httpstco08jjuq1mpa	1	1

Table 6: Sample Tweets and Respective Coding

Table 7: Linear Regression Results for H1

	(1)	(2)	(3)	(4)
Predicted Ideology:	D	R	D	R
Is the Primary Now? (1 = yes)	-0.0259 (-7.82)	-0.00626 (-2.02)	-0.0257 (-7.78)	0.00711 (2.33)
Word Count			-0.0132 (-18.82)	-0.0200 (-25.85)
Character Count			0.00266 (32.43)	0.00417 (43.29)
Average Word Length			-0.00227 (-2.09)	-0.00476 (-4.84)
Number of Stop Words			-0.00577 (-7.41)	-0.00582 (-7.31)
Number of Upper Case Characters			0.0283 (19.43)	0.0354 (29.00)
Number of Hashtags				0.0108 (8.15)
Number of Numerics				-0.0278 (-6.81)
Constant	1.381 (532.12)	1.448 (601.77)	1.303 (145.61)	1.283 (150.81)
Observations	127363	135417	127363	135417

t statistics in parentheses

Table 8: Fixed Effects Regression Results for H1

	(1)	(2)	(3)	(4)
Predicted Ideology:	(mean) D	(mean) R	(mean) D	(mean) R
Is the Primary Now? (1 = yes)	-0.00886 (-1.45)	0.0104 (1.91)	-0.0149 (-2.47)	0.00503 (0.94)
(mean) Word Count			-0.0180 (-10.12)	-0.0179 (-10.34)
(mean) Character Count			0.00334 (15.01)	0.00422 (19.36)
(mean) Number of Stop Words			-0.00573 (-2.83)	-0.0102 (-5.58)
(mean) Number of Upper Case Characters			0.0366 (9.15)	0.0395 (14.09)
(mean) Average Word Length				-0.00653 (-3.29)
(mean) Number of Hashtags				0.00752 (2.41)
(mean) Number of Numerics				-0.0178 (-2.03)
Constant	1.363 (289.11)	1.433 (336.96)	1.274 (95.18)	1.267 (67.37)
Observations	20565	27507	20565	27507

t statistics in parentheses

H2

As H1 was proven significant through the controlled linear regressions, albeit weakly, H2 remains an interesting and viable hypothesis. If the logic of H1 holds true, we should observe a greater negative coefficient for Democrats in states with closed primary rules when compared to open primary rules, and a greater positive coefficient for Republicans in states with closed primary rules when compared to open primary rules. Such a result accurately describes the idea that Democrats and Republicans will be more solidified within their bases during the primary season if the primary is closed versus open. Table 10 below contains the results from the bi-variate regression.

While H2 appears to hold true for Democrats, it would appear Republicans are left of their General Election positions in closed primaries, which makes no sense. Furthermore, the results for open Republican primaries are not significant. Table 11 below contains results for the multi-variate regressions with the non-significant controls omitted. Even here, however, we see the same trend for the Republican Party in closed primaries.

This could be a result of omitted variable bias and therefore warrant the use of either fixed- or random-effects. The Hausman specification test supports the use of random-effects model for all four bi-variate regressions, the results of which are displayed in Table 12, where we observe almost universally insignificant results. Multi-variate regressions controlling for tweet attributes can be found in Table 13. The Hausman specification test determined the following for each model: Models 1 and 3 are best suited for a random-effects treatment, and Models 2 and 4 for a fixed-effects treatment. We find statistically insignificant results across the board here as well.

We must then assume that the results as shown in Table 11 are true, and that Republicans observably move to the left during closed primaries when compared to their position in the general. Therefore, H2 is proven for Democrats, but inconclusive for Republicans due to the null result for the fourth model. This so strongly counters intuition that it seems more likely that there is an issue with the input data although there is none that I can detect.

Table 9: Linear Regression Results for H2

	(1)	(2)	(3)	(4)
Predicted Ideology:	D, Closed	D, Open	R, Closed	R, Open
Is the Primary Now? (1 = yes)	-0.0415 (-6.01)	-0.0186 (-4.72)	-0.0416 (-6.88)	0.00606 (1.59)
Constant	1.345 (245.12)	1.393 (446.56)	1.439 (324.54)	1.464 (482.80)
Observations	32221	86843	35823	88993

t statistics in parentheses

Table 10: Linear Regression Results With Controls for H2

	(1)	(2)	(3)	(4)
Predicted Ideology:	D, Closed	D, Open	D, Closed	D, Open
Is the Primary Now? (1 = yes)	-0.0451 (-6.48)	-0.0154 (-3.91)	-0.0315 (-5.23)	0.0223 (5.95)
Word Count	-0.0177 (-14.55)	-0.0123 (-17.20)	-0.0220 (-13.34)	-0.0197 (-21.88)
Character Count	0.00293 (15.05)	0.00263 (32.79)	0.00389 (18.98)	0.00427 (38.34)
Average Word Length	-0.0103 (-5.01)		-0.00862 (-4.02)	-0.00589 (-5.15)
Number of Hashtags	-0.0253 (-7.20)	0.0126 (6.81)		0.0135 (8.85)
Number of Numerics	-0.0194 (-2.42)		-0.0202 (-2.62)	-0.0293 (-5.87)
Number of Upper Case Characters	0.0380 (11.40)	0.0232 (14.03)	0.0427 (17.19)	0.0314 (21.87)
Number of Stop Words		-0.00645 (-7.03)	-0.00342 (-2.11)	-0.00753 (-7.95)
Constant	1.355 (74.47)	1.285 (241.94)	1.374 (71.91)	1.288 (128.02)
Observations	32221	86843	35823	88993

t statistics in parentheses

Table 11: Fixed- and Random-Effects Regression Results for H2

	(1)	(2)	(3)	(4)
Predicted Ideology:	D, Closed	D, Open	D, Closed	D, Open
Is the Primary Now? (1 = yes)	0.000950 (0.09)	-0.0109 (-1.40)	-0.0129 (-1.23)	0.0216 (3.21)
Constant	1.354 (87.88)	1.376 (110.50)	1.434 (86.41)	1.447 (114.74)
Observations	7292	11964	7237	17672

t statistics in parentheses

Table 12: Fixed- and Random-Effects Regression Results With Controls for H2

	(1)	(2)	(3)	(4)
Predicted Ideology (mean):	D, Closed	D, Open	D, Closed	D, Open
Is the Primary Now? (1 = yes)	-0.0101 (-0.96)	-0.0145 (-1.87)	-0.0129 (-1.26)	0.0112 (1.68)
(mean) Word Count	-0.0237 (-11.24)	-0.0155 (-6.96)	-0.0158 (-5.31)	-0.0179 (-8.42)
(mean) Character Count	0.00373 (10.24)	0.00318 (11.47)	0.00420 (10.95)	0.00408 (15.38)
(mean) Number of Upper Case Characters	0.0521 (7.57)	0.0270 (5.34)	0.0347 (6.70)	0.0407 (11.68)
(mean) Number of Stop Words		-0.00854 (-3.28)	-0.0129 (-3.63)	-0.00962 (-4.32)
(mean) Average Word Length				-0.00820 (-3.57)
(mean) Number of Hashtags				0.0131 (3.44)
(mean) Number of Numerics				-0.0219 (-2.00)
Constant	1.287 (51.23)	1.277 (75.28)	1.223 (49.00)	1.298 (58.75)
Observations	7292	11964	7237	17672

t statistics in parentheses

Summary of Results

This investigation started as something grounded in the traditional literature surrounding the median voter theorem and spatial elections. Beginning with Downs (1957), the median voter theorem as dominated the academic landscape of polarization, candidate positioning, and elections. Since then, countless papers and books have been published supporting the theory put forth by Downs from both empirical and theoretical angles. However, few to date have looked to political communication as a potential avenue for investigating the veracity of the claims of the median voter theorem in action as seen in the American primary system. I set out to prove its truth through a modern framework that bridged data science and political and economic theory, connecting real-world data and outcomes to our best predictions of what should happen given rational voting populations and politicians. Based on the literature, I formed two hypotheses: the first predicts that Republicans and Democrats will pursue their party's median voter during the primary elections and then pivot to the country's true median voter during the general, and the second that this effect would be larger in states with closed primaries versus open primaries (where the median voter could ostensibly be closer to the country's political center).

H1 has been proven empirically in this paper. I justified the hypothesis by citing the seminal papers on the median voter theorem, including Holcombe (1980),¹²⁴ Rice (1985),¹²⁵ Husted and Kenny (1997)¹²⁶ Fujiwara (2015).¹²⁷ The linear regression models tested in pursuit of H1 have shown that, although the effects observed are weak, they are still

124. HOLCOMBE, "AN EMPIRICAL TEST OF THE MEDIAN VOTER MODEL."

125. Rice, "An Examination of the Median Voter Hypothesis."

126. Husted and Kenny, "The effect of the expansion of the voting franchise on the size of government."

127. Fujiwara, "Voting Technology, Political Responsiveness, and Infant Health: Evidence From Brazil."

statistically significant and in the expected direction for each party. Controlling for tweet attributes the effects get stronger. This proof is both supported by and supports the assertions made by Jacobsen (2015)¹²⁸ and Sinclair (2006)¹²⁹ which predict polarized electorates when comparing the voting population of primaries to general elections. Assuming that politicians have a reasonably accurate knowledge of their electorate, it stands to reason that they would moderate expressed political ideology in a fashion collinear with that of the true political preferences of eligible voters during that period of the election. Further, fixed-effects analysis of H1 showed us that while this trend is observable at a candidate-level in the Democratic Party, the same is not true for Republicans. It is unclear why this is true, but potential issues include omitted variable bias or heterogeneity in the amount or quality of data that could be collected for each candidate. Later analysis could be directed to validate and control for this.

H2, however, yielded a more complex result in that it can only be proven for Democrats. Multi-variate regression analysis revealed the expected pattern among Democrats but was inconclusive concerning Republicans. My justification of H2 was simply that if H1 were to hold true, then H2 would have to be true assuming politicians know their electorate. Because the results are inconclusive, however, it is difficult to opine on potential reasons why this is past weaknesses in the data.

This paper also supports the use of data science approaches to political science through its use of Twitter, supervised machine learning, and natural language processing. Although Twitter is used as an investigative tool in formal political science extremely

128. Jacobson, *The politics of congressional elections*.

129. Sinclair, *Party wars: polarization and the politics of national policy making*.

widely outside of the United States, academia in America has been surprisingly conservative in adopting its use outside of the realm of political communication. The increasing importance of Twitter to the modern political landscape, as justified by LaMarre (2013)¹³⁰ and Evans et al. (2014),¹³¹ and similar studies around the world, almost demands its adoption into the discipline. Twitter is, as seen here, an extremely useful and portable source of massive amounts of information.

Although I never aimed to answer the question of whether or not primaries are causing polarization, I did want to look at the connection between primaries and the ideology of political communication as a potential avenue for future investigation in this area. It would seem from the results that this is an avenue with potential fruit, some of which I have uncovered but much of which is still open to the political science community to pursue. The results found here fit within the context of the dominant literature such as McKelvey and Ordeshook's experiments, Bradey et al. (2007), Burden (2004), and the literature cited above while also fitting within the framework of Hirano et al. (2010), McGhee et al. (2014), and their lineage. I don't challenge the assertions that primaries don't lead to a polarized result in elections, but prior to this paper, the question of how primaries affect electoral dynamics was still an unanswered one. I see this paper as a step in the right direction towards answering it.

130. LaMarre and Suzuki-Lambrecht, "Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns'."

131. Evans, Cordova, and Sipole, "Twitter style: An analysis of how house candidates used twitter in their 2012 campaigns."

Chapter 4

While I have found some significant results in this paper, there are still significant weaknesses with regards to data availability and analysis, and competing literature that warrant mention and future study. There are also many areas for improvement that could be targeted in a revision of this paper but were not possible here given the time and resource limitations for this submission.

The first major weakness of this paper is the data, under which several smaller weaknesses contribute. I will split this into two categories: data availability and data analysis.

Data availability for Twitter accounts was not perfect. As indicated by Table 2, only 37.61% of the candidates running for Senate across the three observable election cycles had Twitter accounts that could be parsed for data. Although a total of 262,780 tweets were collected, that is an average of 1,030.5 tweets per candidate in a dataset that saw some candidates tweet as few as 7 times over the course of the election, and others over 20,000 times. This could lead to biased results. Although I tried to accommodate for this by excluding outlandishly prolific tweeters from the dataset, this is still a potential issue. Within this, there is also a concern about unobserved traits that could cause someone's tweet volume to be more or less, including candidate-level characteristics, race-level characteristics, or geographic characteristics. There has been literature that shows deviation in tweeting due to candidate-level characteristics, including Valenzuela et al.

(2009),¹³² Evans et al. (2014),¹³³ McGregor et al. (2016),¹³⁴ Evans et al. (2016),¹³⁵ McGregor et al. (2018),¹³⁶ and Evans et al. (2018),¹³⁷ and Burden (2004)¹³⁸ cites race competition as a potential interaction factor with expressed ideology. Geographic characteristics such as age of the voting electorate, internet availability, and the veracity of retail politics are all potential reasons for a candidate NOT to use Twitter to campaign, and things that the dataset does not control for.

Data analysis is also a potential weakness, especially when it comes to the supervised learning approach. Algorithms are far from perfect and there is a massive potential for errors in the coding that could affect the results seen in this paper. That the results proven here are in line with one would expect are encouraging and bolster the potential of supervised machine learning to make its way into political science research. This concern should not be taken as a reason to discredit the results found here. However, it is possible that the reason the effects found are so weak, and not as strongly statistically significant as one would expect, are because of a weakness with the training set when compared to the input data, a potential weakness of the approach itself when it comes to

132. Sebastián Valenzuela and Teresa Correa, “Press Coverage and Public Opinion On Women Candidates,” *International Communication Gazette* 71, no. 3 (April 2009): 203–223, ISSN: 1748-0485, doi:10.1177/1748048508100913, <http://journals.sagepub.com/doi/10.1177/1748048508100913>.

133. Evans, Cordova, and Sipole, “Twitter style: An analysis of how house candidates used twitter in their 2012 campaigns.”

134. Shannon C. McGregor and Rachel R. Mourão, “Talking Politics on Twitter: Gender, Elections, and Social Networks,” *Social Media + Society* 2, no. 3 (September 2016): 205630511666421, ISSN: 2056-3051, doi:10.1177/2056305116664218, <http://journals.sagepub.com/doi/10.1177/2056305116664218>.

135. Heather K. Evans and Jennifer Hayes Clark, ““You Tweet Like a Girl!”,” *American Politics Research* 44, no. 2 (March 2016): 326–352, ISSN: 1532-673X, doi:10.1177/1532673X15597747, <http://journals.sagepub.com/doi/10.1177/1532673X15597747>.

136. Shannon C McGregor, “Personalization, social media, and voting: Effects of candidate self-personalization on vote intention,” *New Media & Society* 20, no. 3 (March 2018): 1139–1160, ISSN: 1461-4448, doi:10.1177/1461444816686103, <http://journals.sagepub.com/doi/10.1177/1461444816686103>.

137. Heather K. Evans, Kayla J. Brown, and Tiffany Wimberly, ““Delete Your Account”,” *Social Science Computer Review* 36, no. 4 (August 2018): 500–508, ISSN: 0894-4393, doi:10.1177/0894439317728722, <http://journals.sagepub.com/doi/10.1177/0894439317728722>.

138. Burden, *Candidate Positioning in US Congressional Elections*.

this study.

A major concern going into this project was that the training set used for the supervised learning algorithms was trained on presidential candidate tweets, whereas the inputs were tweets from candidates for Senate. It is not known whether there are significant differences in the way that candidates for Senate tweet when compared to candidates for President, but it is definitely possible that there are additional flags that went unobserved by the algorithm that are unique to Senate races and may have predicted a different categorization for a given tweet. Of further concern is the unknown amount of heterogeneity between candidates from different states; it may be entirely inappropriate to use the same algorithm to sort tweets of candidates from different states, and although there is no evidence that this is the case, that is only because this is yet unknown.

Another related concern has to do with the method of supervised machine learning itself. Text classification has traditionally been performed on larger documents with more features. At base, a tweet is a maximum of 140 words, and after necessary text cleaning, the the document generally shortens. This means that the number of relevant features goes down, potentially decreasing the accuracy of the model. I am not aware of any particular models that have been optimized for classifying short documents in particular, but advanced supervised learning models such as BERT show promise for future work in natural language processing, albeit at a high processing cost. This is a significant drawback of the approach that is hard to validate, and although it is unlikely that it would bias the results found in this paper, it is possible that it has interfered with the degree to which we would expect to see the results we found.

Finally, competing literature such as Serra (2015) and (2018) are not defeated by

the results here, nor do the results found there nullify that which is proven here. Rather, when one gap is filled by this paper, another opens up due to the incongruity of the theoretical and empirical results. Open still is the question of what causes the phenomenon proven in Chapter 3. The most that can be concluded from this paper is that the presence of primaries, and particular types of primaries, has particular effects on candidate ideology in expected directions for each party. However, it has not been conclusively shown here that primaries are the only factor responsible for this phenomenon. Serra left open the potential for some other factor that, through the framework of the primary system, cause polarized candidates to enter the general election. Is it possible that this unknown factor is also present here, affecting candidate ideology in a manner that is consistent with theoretical literature about the median voter theorem through the framework of primaries, but detached from primary elections itself? This is an interesting question for future study that could bridge this critical gap in the research.

References

- Acree, Brice D.L., Justin H. Gross, Noah A. Smith, Yanchuan Sim, and Amber E. Boydstun. "Etch-a-Sketching: Evaluating the Post-Primary Rhetorical Moderation Hypothesis." *American Politics Research* 48, no. 1 (January 2020): 99–131. ISSN: 1532-673X. doi:10.1177/1532673X18800017. <http://journals.sagepub.com/doi/10.1177/1532673X18800017>.
- Ahmad, Akintunde, Adam Gabbatt, Lauren Harris, Savannah Jacobson, Ed Pilkington, and Jim Waterson. *Sleepwalking into 2020 - Columbia Journalism Review*, 2020. Accessed February 22, 2020. https://www.cjr.org/special%7B%5C_%7Dreport/2020-election-media-journalism.php.
- ANDERSON, PETER J., and AILEEN McLEOD. "The Great Non-Communicator? The Mass Communication Deficit of the European Parliament and its Press Directorate*." *JCMS: Journal of Common Market Studies* 42, no. 5 (December 2004): 897–917. ISSN: 0021-9886. doi:10.1111/j.0021-9886.2004.00534.x. <http://doi.wiley.com/10.1111/j.0021-9886.2004.00534.x>.
- Andrew Dugan. *Majority of Americans See Congress as Out of Touch, Corrupt*, 2015. Accessed February 8, 2020. <https://news.gallup.com/poll/185918/majority-americans-congress-touch-corrupt.aspx>.
- Ansolabehere, Stephen, and James M. Snyder. "The Incumbency Advantage in U.S. Elections: An Analysis of State and Federal Offices, 1942–2000." *Election Law Journal: Rules, Politics, and Policy* 1, no. 3 (September 2002): 315–338. ISSN: 1533-1296.

doi:10.1089/153312902760137578.

<http://www.liebertpub.com/doi/10.1089/153312902760137578>.

Ashworth, Scott, and Ethan Bueno De Mesquita. "Electoral selection, strategic challenger entry, and the incumbency advantage." *Journal of Politics*, 2008. ISSN: 00223816.

doi:10.1017/S0022381608081024.

Bartels, Larry M. *Opinion / The Rise of Presidential Extremists - The New York Times*, 2016. Accessed December 7, 2019.

<https://www.nytimes.com/2016/09/12/opinion/campaign-stops/the-rise-of-presidential-extemists.html>.

Bekafigo, Marija, and Allison Clark Pingley. "Do Campaigns "Go Negative" on Twitter?"

In *Politics, Protest, and Empowerment in Digital Spaces*, 178–191. IGI Global,

December 2017. ISBN: 9781522518631. doi:10.4018/978-1-5225-1862-4.ch011.

<http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-1862-4.ch011>.

Benton, J. Edwin. *When Americans go to the polls, they look to the past – not the future*,

2019. Accessed February 25, 2020. <https://theconversation.com/when-americans-go-to-the-polls-they-look-to-the-past-not-the-future-114402>.

Blais, Andre, Elisabeth Gidengil, and Neil Nevitte. "Where does turnout decline come

from?" *European Journal of Political Research* 43, no. 2 (March 2004): 221–236. ISSN: 0304-4130. doi:10.1111/j.1475-6765.2004.00152.x.

<http://doi.wiley.com/10.1111/j.1475-6765.2004.00152.x>.

Blais, André, and Daniel Rubenson. "The Source of Turnout Decline." *Comparative*

Political Studies 46, no. 1 (January 2013): 95–117. ISSN: 0010-4140.

doi:10.1177/0010414012453032.

<http://journals.sagepub.com/doi/10.1177/0010414012453032>.

Bode, Leticia, Alexander Hanna, Junghwan Yang, and Dhavan V. Shah. "Candidate

Networks, Citizen Clusters, and Political Expression." *The ANNALS of the American*

Academy of Political and Social Science 659, no. 1 (May 2015): 149–165. ISSN:

0002-7162. doi:10.1177/0002716214563923.

<http://journals.sagepub.com/doi/10.1177/0002716214563923>.

Boots, Ralph S. "American Parties and Elections. BY Edward M. Sait. (New York: The

Century Company. 1927. Pp. 608.)" *American Political Science Review* 21, no. 4

(September 1927): 900–902. ISSN: 0003-0554. doi:10.2307/1947614.

[https://www.cambridge.org/core/product/identifier/S0003055400024989/
type/journal%7B%5C_%7Darticle](https://www.cambridge.org/core/product/identifier/S0003055400024989/type/journal%7B%5C_%7Darticle).

Bossetta, Michael. "The Digital Architectures of Social Media: Comparing Political

Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S.

Election." *Journalism & Mass Communication Quarterly* 95, no. 2 (June 2018):

471–496. ISSN: 1077-6990. doi:10.1177/1077699018763307.

<http://journals.sagepub.com/doi/10.1177/1077699018763307>.

Bowman, Ann OM., and Richard C. Kearney. *State and local government: the essentials*.

Cengage Learning, 2015.

- Brady, David W., Hahrie Han, and Jeremy C. Pope. "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?" *Legislative Studies Quarterly* 32, no. 1 (February 2007): 79–105. ISSN: 03629805. doi:10.3162/036298007X201994.
<http://doi.wiley.com/10.3162/036298007X201994>.
- Broockman, David E, and Christopher Skovron. *What Politicians Believe About Their Constituents: Asymmetric Misperceptions and Prospects for Constituency Control* *. Technical report. 2013. <http://polisci.berkeley.edu/people/graduatestudents/persondetail.php?person=381>.<http://sitemaker.umich.edu/cskovron/home.1>.
- Bryden, John, and Eric Silverman. "Underlying socio-political processes behind the 2016 US election." Edited by Haroldo V. Ribeiro. *PLOS ONE* 14, no. 4 (April 2019): e0214854. ISSN: 1932-6203. doi:10.1371/journal.pone.0214854.
<http://dx.plos.org/10.1371/journal.pone.0214854>.
- Burden, Barry C. *Candidate Positioning in US Congressional Elections*, 2004.
doi:10.1017/S000712340400002X.
- Campbell, James E., Bryan J. Dettrey, and Hongxing Yin. "The theory of conditional retrospective voting: Does the presidential record matter less in open-seat elections?" *Journal of Politics* 72, no. 4 (October 2010): 1083–1095. ISSN: 00223816.
doi:10.1017/S002238161000054X.
- Carson, Jamie L., and Ryan D. Williamson. "Candidate ideology and electoral success in congressional elections." *Public Choice* 176, nos. 1-2 (July 2018): 175–192. ISSN: 15737101. doi:10.1007/s11127-017-0492-2.

Ceron, Andrea, and Giovanna D'Adda. "E-campaigning on Twitter: The effectiveness of distributive promises and negative campaign in the 2013 Italian election." *New Media & Society* 18, no. 9 (October 2016): 1935–1955. ISSN: 1461-4448.

doi:10.1177/1461444815571915.

<http://journals.sagepub.com/doi/10.1177/1461444815571915>.

Cho, Wendy K. Tam. *Voter migration is a significant factor in the geographic sorting of the American electorate* / USAPP, 2018. Accessed December 9, 2019.

<https://blogs.lse.ac.uk/usappblog/2013/11/19/voter-migration-is-a-significant-factor-in-the-geographic-sorting-of-the-american-electorate/>.

Chong, Miyoung. "Analyzing political information network of the U.S. Partisan public on twitter." In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10766 LNCS, 453–463. Springer Verlag, 2018. ISBN: 9783319781044. doi:10.1007/978-3-319-78105-1_50.

Cohen, Marty, Mary C. McGrath, Peter Aronow, and John Zaller. "Ideologically Extreme Candidates in U.S. Presidential Elections, 1948–2012." Edited by Larry M. Bartels. *The ANNALS of the American Academy of Political and Social Science* 667, no. 1 (September 2016): 126–142. ISSN: 0002-7162. doi:10.1177/0002716216660571.

<http://journals.sagepub.com/doi/10.1177/0002716216660571>.

Collier, Kenneth E., Richard D. McKelvey, Peter C. Ordeshook, and Kenneth C. Williams. "Retrospective voting: An experimental study." *Public Choice* 53, no. 2 (January 1987): 101–130. ISSN: 00485829. doi:10.1007/BF00125844.

Collier, Kenneth, Peter C. Ordeshook, and Kenneth Williams. "The rationally uninformed electorate: Some experimental evidence." *Public Choice* 60, no. 1 (January 1989): 3–29. ISSN: 00485829. doi:10.1007/BF00124309.

Conkling, A.R. "Reviews : Nominating Systems: Direct Primaries versus Conventions in the United States. By ERNST C. MEYER. Pp. 50I. Price, \$I.50. Published by the Author. Madison, Wisconsin, 1902." *The ANNALS of the American Academy of Political and Social Science* 21, no. 1 (January 1903): 112–114. ISSN: 0002-7162. doi:10.1177/000271620302100111. <http://journals.sagepub.com/doi/10.1177/000271620302100111>.

Dallinger, Frederick William. *Nominations for Elective Office in the United States*. Vol. 4. Longmans, Green, / Company, 1897.

Darius, Philipp, and Fabian Stephany. "'Hashjacking' the Debate: Polarisation Strategies of Germany's Political Far-Right on Twitter." In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11864 LNCS, 298–308. Springer, November 2019. ISBN: 9783030349707. doi:10.1007/978-3-030-34971-4_21.

Djupe, Paul A., and David A. M. Peterson. "The Impact of Negative Campaigning: Evidence from the 1998 Senatorial Primaries." *Political Research Quarterly* 55, no. 4 (December 2002): 845–860. ISSN: 1065-9129. doi:10.1177/106591290205500406. <http://journals.sagepub.com/doi/10.1177/106591290205500406>.

Downs, Anthony. "An Economic Theory of Political Action in a Democracy." *Journal of Political Economy*, 1957. ISSN: 0022-3808. doi:10.1086/257897.

Ernst, Nicole, Sina Blassnig, Sven Engesser, Florin Büchel, and Frank Esser. "Populists Prefer Social Media Over Talk Shows: An Analysis of Populist Messages and Stylistic Elements Across Six Countries." *Social Media + Society* 5, no. 1 (January 2019): 205630511882335. ISSN: 2056-3051. doi:10.1177/2056305118823358.
<http://journals.sagepub.com/doi/10.1177/2056305118823358>.

Evans, Heather K., Kayla J. Brown, and Tiffany Wimberly. "'Delete Your Account'." *Social Science Computer Review* 36, no. 4 (August 2018): 500–508. ISSN: 0894-4393. doi:10.1177/0894439317728722.
<http://journals.sagepub.com/doi/10.1177/0894439317728722>.

Evans, Heather K., and Jennifer Hayes Clark. "'You Tweet Like a Girl!'" *American Politics Research* 44, no. 2 (March 2016): 326–352. ISSN: 1532-673X. doi:10.1177/1532673X15597747.
<http://journals.sagepub.com/doi/10.1177/1532673X15597747>.

Evans, Heather K., Victoria Cordova, and Savannah Sipole. "Twitter style: An analysis of how house candidates used twitter in their 2012 campaigns." *PS - Political Science and Politics* 47, no. 2 (2014): 454–462. ISSN: 15375935. doi:10.1017/S1049096514000389.

Francia, Peter L. "Free Media and Twitter in the 2016 Presidential Election." *Social Science Computer Review* 36, no. 4 (August 2018): 440–455. ISSN: 0894-4393.

doi:10.1177/0894439317730302.

<http://journals.sagepub.com/doi/10.1177/0894439317730302>.

Fridkin, Kim Leslie, and Patrick J. Kenney. "Do Negative Messages Work?" *American Politics Research* 32, no. 5 (September 2004): 570–605. ISSN: 1532-673X.

doi:10.1177/1532673X03260834.

<http://journals.sagepub.com/doi/10.1177/1532673X03260834>.

Fujiwara, Thomas. "Voting Technology, Political Responsiveness, and Infant Health: Evidence From Brazil." *Econometrica* 83, no. 2 (March 2015): 423–464. ISSN: 0012-9682. doi:10.3982/ecta11520.

Gervais, Bryan T., and Walter C. Wilson. "New media for the new electorate? Congressional outreach to Latinos on Twitter." *Politics, Groups, and Identities* 7, no. 2 (April 2019): 305–323. ISSN: 2156-5503. doi:10.1080/21565503.2017.1358186.
<https://www.tandfonline.com/doi/full/10.1080/21565503.2017.1358186>.

Griffin, Hannah. "Keep it Clean? How Negative Campaigns Affect Voter Turnout." *Res Publica - Journal of Undergraduate Research* 17, no. 1 (July 2012).
<https://digitalcommons.iwu.edu/respublica/vol17/iss1/6>.

Gross, Justin H., and Kaylee T. Johnson. "Twitter Taunts and Tirades: Negative Campaigning in the Age of Trump." *PS: Political Science & Politics* 49, no. 04 (October 2016): 748–754. ISSN: 1049-0965. doi:10.1017/S1049096516001700.
http://www.journals.cambridge.org/abstract%7B%5C_%7DS1049096516001700.

Gruber, Jonathan. *Public Finance and Public Policy* (3rd). 860. Worth Publishers, 2011.

ISBN: 9781429219495.

Hemphill, Libby, and Matthew A. Shapiro. “Appealing to the base or to the moveable middle? Incumbents’ partisan messaging before the 2016 U.S. congressional elections.”

Journal of Information Technology & Politics 16, no. 4 (October 2019): 325–341. ISSN: 1933-1681. doi:10.1080/19331681.2019.1651685.

<https://www.tandfonline.com/doi/full/10.1080/19331681.2019.1651685>.

Henderson, C. R. “Anthracite Coal Communities. Peter Roberts.” *American Journal of Sociology* 9, no. 6 (May 1904): 844–845. ISSN: 0002-9602. doi:10.1086/211277.

<https://www.journals.uchicago.edu/doi/10.1086/211277>.

Hershey, Marjorie Randon, and John H. Aldrich. *Party politics in America*. 2017. ISBN: 9781134836666. doi:10.4324/9781315544427.

Hertel-Fernandez, Alexander, Matto Mildenberger, and Leah C. Stokes. *Opinion / Congress Has No Clue What Americans Want - The New York Times*, 2018. Accessed February 8, 2020. <https://www.nytimes.com/2018/10/31/opinion/congress-midterms-public-opinion.html>.

Hirano, Shigeo, and James M. Snyder, Jr. *Primary Elections in the United States*.

Cambridge University Press, September 2019. ISBN: 9781139946537.

doi:10.1017/9781139946537. [https:](https://www.cambridge.org/core/product/identifier/9781139946537/type/book)

[//www.cambridge.org/core/product/identifier/9781139946537/type/book](https://www.cambridge.org/core/product/identifier/9781139946537/type/book).

Hirano, Shigeo, James M. Snyder, Stephen Ansolabehere, and John Mark Hansen.

“Primary elections and partisan polarization in the U.S. Congress.” *Quarterly Journal of Political Science* 5, no. 2 (2010): 169–191. ISSN: 15540626.

doi:10.1561/100.00008052.

HOLCOMBE, RANDALL G. “AN EMPIRICAL TEST OF THE MEDIAN VOTER

MODEL.” *Economic Inquiry* 18, no. 2 (April 1980): 260–274. ISSN: 14657295.

doi:10.1111/j.1465-7295.1980.tb00574.x.

<http://doi.wiley.com/10.1111/j.1465-7295.1980.tb00574.x>.

Holcombe, Randall G. *Public sector economics : the role of government in the American economy*. 517. Pearson Education, 2006. ISBN: 9780131450424.

Hopkins, Daniel J., and Lindsay M. Pettingill. “Retrospective Voting in Big-City US

Mayoral Elections.” *Political Science Research and Methods* 6, no. 4 (October 2018):

697–714. ISSN: 20498489. doi:10.1017/psrm.2016.54.

Husted, Thomas A., and Lawrence W. Kenny. “The effect of the expansion of the voting

franchise on the size of government.” *Journal of Political Economy* 105, no. 1 (October

1997): 54–82. ISSN: 00223808. doi:10.1086/262065.

Isaac, Mike, and Sydney Ember. “For Election Day Influence, Twitter Ruled Social Media.”

<https://www.nytimes.com/2016/11/09/technology/for-election-day-chatter-twitter-ruled-social-media.html>. (accessed: 06.06.2019).

Jacobson, Gary C. *The politics of congressional elections*. 321. Rowman & Littlefield, 2015.

ISBN: 1442252634.

- Kahn, Kim Fridkin, and Patrick J. Kenney. "Do Negative Campaigns Mobilize or Suppress Turnout? Clarifying the Relationship between Negativity and Participation." *American Political Science Review* 93, no. 4 (December 1999): 877–889. ISSN: 0003-0554.
doi:10.2307/2586118. https://www.cambridge.org/core/product/identifier/S000305540021914X/type/journal%7B%5C_%7Darticle.
- Keisling, Phil. *To Reduce Partisanship, Get Rid of Partisans - The New York Times*, 2010.
Accessed December 8, 2019.
<https://www.nytimes.com/2010/03/22/opinion/22keisling.html>.
- Keller, Franziska B., David Schoch, Sebastian Stier, and JungHwan Yang. "Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign." *Political Communication*, October 2019, 1–25. ISSN: 1058-4609.
doi:10.1080/10584609.2019.1661888.
<https://www.tandfonline.com/doi/full/10.1080/10584609.2019.1661888>.
- LaMarre, Heather L., and Yoshikazu Suzuki-Lambrecht. "Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns'." *Public Relations Review* 39, no. 4 (November 2013): 360–368. ISSN: 03638111.
doi:10.1016/j.pubrev.2013.07.009.
- Lanoue, David J. "Retrospective and Prospective Voting in Presidential-Year Elections." *Political Research Quarterly* 47, no. 1 (March 1994): 193. ISSN: 10659129.
doi:10.2307/448908.

McCarthy, Justin. *Trump Seen as Less Conservative Than Prior GOP Candidates*.

Accessed December 7, 2019. <https://news.gallup.com/poll/196064/trump-seen-less-conservative-prior-gop-candidates.aspx>.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. "The Polarization of the Politicians." In *Polarized America: The Dance of Ideology and Unequal Riches*. 2006.

McGhee, Eric, Seth Masket, Boris Shor, Steven Rogers, and Nolan McCarty. "A Primary Cause of Partisanship? Nomination Systems and Legislator Ideology." *American Journal of Political Science*, 2014. ISSN: 15405907. doi:10.1111/ajps.12070.

McGregor, Shannon C. "Personalization, social media, and voting: Effects of candidate self-personalization on vote intention." *New Media & Society* 20, no. 3 (March 2018): 1139–1160. ISSN: 1461-4448. doi:10.1177/1461444816686103. <http://journals.sagepub.com/doi/10.1177/1461444816686103>.

McGregor, Shannon C., and Rachel R. Mourão. "Talking Politics on Twitter: Gender, Elections, and Social Networks." *Social Media + Society* 2, no. 3 (September 2016): 205630511666421. ISSN: 2056-3051. doi:10.1177/2056305116664218. <http://journals.sagepub.com/doi/10.1177/2056305116664218>.

McKelvey, Richard D., and Peter C. Ordeshook. "Elections with limited information: A fulfilled expectations model using contemporaneous poll and endorsement data as information sources." *Journal of Economic Theory* 36, no. 1 (June 1985): 55–85. ISSN: 10957235. doi:10.1016/0022-0531(85)90079-1.

McKelvey, Richard D., and Peter C. Ordeshook. "Elections with limited information: A multidimensional model." *Mathematical Social Sciences* 14, no. 1 (August 1987):

77–99. ISSN: 01654896. doi:10.1016/0165-4896(87)90016-3.

———. "Rational expectations in elections: some experimental results based on a multidimensional model." *Public Choice* 44, no. 1 (January 1984): 61–102. ISSN: 00485829. doi:10.1007/BF00124819.

———. "Sequential Elections with Limited Information." *American Journal of Political Science* 29, no. 3 (August 1985): 480. ISSN: 00925853. doi:10.2307/2111140.

———. "Two-Candidate Elections without Majority Rule Equilibria." *Simulation & Games* 13, no. 3 (September 1982): 311–335. ISSN: 0037-5500.

doi:10.1177/0037550082133004.

<http://journals.sagepub.com/doi/10.1177/0037550082133004>.

Mclaren, Craig B. *DOES GERRYMANDERING VIOLATE THE FOURTEENTH AMENDMENT?: INSIGHT FROM THE MEDIAN VOTER THEOREM*. Technical report. 2017.

Min, Young. "News Coverage of Negative Political Campaigns." *Harvard International Journal of Press/Politics* 9, no. 4 (October 2004): 95–111. ISSN: 1081-180X.

doi:10.1177/1081180X04271861.

<http://journals.sagepub.com/doi/10.1177/1081180X04271861>.

Morris, David S. "Twitter Versus the Traditional Media." *Social Science Computer Review* 36, no. 4 (August 2018): 456–468. ISSN: 0894-4393. doi:10.1177/0894439317721441.

<http://journals.sagepub.com/doi/10.1177/0894439317721441>.

Pew Research Center. *Public Trust in Government: 1958-2019* / Pew Research Center.

Accessed February 8, 2020. <https://www.people-press.org/2019/04/11/public-trust-in-government-1958-2019/>.

———. *Voters' Perceptions of the Candidates: Traits, Ideology and Impact on Issues* / Pew Research Center, 2016. Accessed December 7, 2019.

<https://www.people-press.org/2016/07/14/voters-perceptions-of-the-candidates-traits-ideology-and-impact-on-issues/>.

Pomper, Gerald M., and Austin Ranney. "Curing the Mischiefs of Faction: Party Reform in America." *Political Science Quarterly*, 1976. ISSN: 00323195. doi:10.2307/2149167.

Presidential Elections: Strategies and Structures of American Politics. 1976.

Reynolds, John F. *The Demise of the American Convention System, 1880–1911*. 2009. doi:10.1017/cbo9780511511684.001.

Rice, T. W. "An Examination of the Median Voter Hypothesis." *Political Research Quarterly* 38, no. 2 (June 1985): 211–223. ISSN: 1065-9129.

doi:10.1177/106591298503800204.

<http://prq.sagepub.com/cgi/doi/10.1177/106591298503800204>.

- Rogowski, Jon C., and Stephanie Langella. "Primary Systems and Candidate Ideology: Evidence From Federal and State Legislative Elections." *American Politics Research*, 2015. ISSN: 15523373. doi:10.1177/1532673X14555177.
- Ross, Andrew S., and David Caldwell. "'Going negative': An APPRAISAL analysis of the rhetoric of Donald Trump on Twitter." *Language and Communication* 70 (January 2020): 13–27. ISSN: 02715309. doi:10.1016/j.langcom.2019.09.003.
- Routledge Handbook of Primary Elections*. 2018. doi:10.4324/9781315544182.
- Sayre, Wallace S., and V. O. Key. "American State Politics: An Introduction." *Political Science Quarterly*, 1956. ISSN: 00323195. doi:10.2307/2145548.
- Schumer, Charles. *Opinion / Charles Schumer: Adopt the Open Primary - The New York Times*, 2014. Accessed December 8, 2019.
<https://www.nytimes.com/2014/07/22/opinion/charles-schumer-adopt-the-open-primary.html>.
- Serra, Gilles. "No Polarization in Spite of Primaries: A Median Voter Theorem with Competitive Nominations," 211–229. 2015. doi:10.1007/978-3-319-15551-7_11.
http://link.springer.com/10.1007/978-3-319-15551-7%7B%5C_%7D11.
- Shapiro, Matthew A., and Libby Hemphill. "Politicians and the Policy Agenda: Does Use of Twitter by the U.S. Congress Direct <i>New York Times</i> Content?" *Policy & Internet* 9, no. 1 (March 2017): 109–132. ISSN: 19442866. doi:10.1002/poi3.120.
<http://doi.wiley.com/10.1002/poi3.120>.

- Shin, Jieun, Lian Jian, Kevin Driscoll, and François Bar. "Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction." *New Media & Society* 19, no. 8 (August 2017): 1214–1235. ISSN: 1461-4448.
doi:10.1177/1461444816634054.
<http://journals.sagepub.com/doi/10.1177/1461444816634054>.
- Sinclair, Barbara. *Party wars: polarization and the politics of national policy making*. 2006.
doi:10.5860/choice.44-2391.
- State Primary Election Types*, 2018. Accessed December 10, 2019.
<http://www.ncsl.org/research/elections-and-campaigns/primary-types.aspx>.
- Stein, Jeff. *Bernie Sanders says Democrats should get rid of closed primaries. Is he right?* - *Vox*, 2016. Accessed December 8, 2019.
<https://www.vox.com/2016/4/28/11469468/open-primaries-closed-primaries-sanders>.
- The Partisan Divide on Political Values Grows Even Wider* / Pew Research Center.
Accessed December 8, 2019. <https://www.people-press.org/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/>.
- Topkis, Jay, and V. O. Key. "Southern Politics in State and Nation." *The Yale Law Journal* 59, no. 6 (May 1950): 1203. ISSN: 00440094. doi:10.2307/793003.
<https://www.jstor.org/stable/793003?origin=crossref>.

Valenzuela, Sebastián, and Teresa Correa. “Press Coverage and Public Opinion On Women Candidates.” *International Communication Gazette* 71, no. 3 (April 2009): 203–223.

ISSN: 1748-0485. doi:10.1177/1748048508100913.

<http://journals.sagepub.com/doi/10.1177/1748048508100913>.

Walter, Annemarie S., and Cees van der Eijk. “Unintended consequences of negative campaigning: Backlash and second-preference boost effects in a multi-party context.”

British Journal of Politics and International Relations 21, no. 3 (August 2019):

612–629. ISSN: 1467856X. doi:10.1177/1369148119842038.

Wojcik, Stefan, and Adam Hughes. “Sizing Up Twitter Users.”

<https://www.pewinternet.org/2019/04/24/sizing-up-twitter-users/>.

(accessed: 06.06.2019).

Woon, Jonathan. “Primaries and Candidate Polarization: Behavioral Theory and

Experimental Evidence.” *American Political Science Review*, 2018. ISSN: 15375943.

doi:10.1017/S0003055418000515.

Yglesias, Matthew. *Moderates win elections: ideological extremism is risky* - *Vox*, 2019.

Accessed December 7, 2019. <https://www.vox.com/2019/7/2/20677656/donald-trump-moderate-extremism-penalty>.

Appendix 1

Measures of Intercode Reliability for Predicted Ideology

	Agreement Rate	Cohen's Kappa
Ideology (Liberal, Neutral, Conservative)	0.78	0.66

Evaluative Metrics for Classification Algorithms (Predicted Ideology)

Name of Algorithm	Input Type	Accuracy	Cohen's Kappa
Naïve Bayes	Count Vectors	0.67	0.00
Naïve Bayes	Word Level TF-IDF	0.69	0.05
Naïve Bayes	N-Gram Vectors	0.67	0.11
Naïve Bayes	CharLevel Vectors	0.68	0.06
Logistical Regression	Count Vectors	0.68	0.00
Logistical Regression	Word Level TF-IDF	0.71	0.20
Logistical Regression	N-Gram Vectors	0.68	0.00
Logistical Regression	CharLevel Vectors	0.71	0.20
Support Vector Machine	N-Gram Vectors	0.69	0.00
Random Forest	Count Vectors	0.68	0.00
Random Forest	Word Level TF-IDF	0.70	0.12
Bagging (Multinomial Bayes)	Count Vectors	0.68	0.12
Bagging (Multinomial Bayes)	Word Level TF-IDF	0.69	0.05
Bagging (Multinomial Bayes)	N-Gram Vectors	0.68	0.01
Bagging (Multinomial Bayes)	CharLevel Vectors	0.69	0.06
Extreme Gradient Boost (XGBoost)	Count Vectors	0.69	0.10
Extreme Gradient Boost (XGBoost)	Word Level TF-IDF	0.68	0.10
Extreme Gradient Boost (XGBoost)	N-Gram Vectors	0.68	0.02
Extreme Gradient Boost (XGBoost)	CharLevel Vectors	0.69	0.11