

Abstract

The 2016 election showed that online, small-dollar donors can impact political campaigns. My research asks: What motivates political donors in Wisconsin state-legislative elections? My analysis finds a link between candidates discussing certain issues online and donations from specific donor communities. In addition, some donor communities are found to be connected by geography. More broadly, this research shows that geography should play a greater role in the study of political communications.

Keywords: social media, political donors, state politics, political geography

It is well-documented that the prevalence of online, small-dollar donors to political campaigns was one of the major stories of the 2016 election cycle (Foran, 2016; Prokop, 2016; Sherfinski, 2017; Stein & Cherkis, 2017; Vogel, 2016). I hypothesized that this change in the culture of political contributions would mean that at least some political donors at all levels of government would be attracted to political campaigns based on the issues that those campaigns talked about on social media. I created a unique methodology that built upon previous scholarly research into the motivations of political donors and attempts to quantify political ideology and predict the issues that political donors care about.

Past research has used political donation information to predict the ideology of donors (Bonica, 2014). But that analysis is limited to predicting ideology in a left-right continuum. I attempt to layer on elite social media data from the candidates receiving political contributions to identify the specific issues that motivate political donors. The output of this study shows that there is a distinct intersection between political donors, online communication, and geography.

Literature Review

Political Donations

Academics have long studied the motivations of political donors. There are three predominant explanations for why donors contribute to campaigns: they seek access and to empower the candidates that are best for their personal interests; they are influenced by their geographic social networks to contribute; or that donors are motivated by policies and geography is largely irrelevant.

The bulk of academic inquiry into political donors and their motivations has focused on access-oriented donors such as large-dollar donors with business interests or political actions committees (PACs). Predominant views of political donations hold that campaign contributions are chiefly investments in a political marketplace “where a return on that investment is expected” (Ansolabehere, de Figueiredo, & Snyder Jr., 2003). Special interests groups have been found to frequently donate to incumbent candidates even if their ideologies do not perfectly align because those groups want to try to attain access to incumbents who are likely to be reelected for a long period of time and attain powerful leadership positions or higher office (Hall, 2014).

Historically, political scientists have seen donors as being primarily access-oriented. But while they may contribute to candidates in both parties, they generally seek to maximize the odds of their preferred party winning, even when making contributions to candidates in their non-preferred party (Brunell, 2005). Overall, case study observations, the opinions of participants and observers, inferences from indirect quantitative analysis, and direct computational analysis have been used as evidence that points to access as an important motivation for campaign contributions. (Langbein, 1986).

More recent studies offer a new paradigm that contributions should be thought less of as investments and more as political consumption (Ansolabehere, de Figueiredo, & Snyder Jr.,

2003). Even individual corporate executives who make contributions that are “often best understood as purchases of ‘good will’ whose returns, while positive in expectation, are contingent and rare (Gordon, Hafer, & Landa, 2007). On the other hand, another study found that 11 percent of CEOs in 2000 who retired by 2011 ended up obtaining political positions—apparent evidence of quid pro quo (Coates IV, 2012). Also, PAC contributions should not be equated to individual donors. “Individuals consistently rank ideological concerns as highly important when deciding where to give” (Barber, 2016).

These ways of seeing political donors was fundamentally changed with the watershed moment of the *Citizens United v. FEC* supreme court ruling. This case ended longstanding campaign finance law limiting political contributions and enabled increased corporate spending to PACs (Kang, 2011). Whether one was personally for or against the ruling, legal and academic researchers largely agreed that it would dramatically change how campaigns were funded and operated (Briffault, 2010; Coates IV, 2012; Epstein, 2011; Gerken, 2014; Hasen, 2011; Kang, 2011). However, there was a subset of academics who questioned whether the change in campaign finance law would have actually have a large practical impact (La Raja & Schaffner, 2014).

In the shadow of *Citizens United*, the growing role of online, small-dollar donors and their potential role in political campaigns was largely overlooked by both scholars and modern political actors. However, recent political campaigns have brought these donations to the forefront of present-day campaigns (Foran, 2016; Prokop, 2016; Sherfinski, 2017; Stein & Cherkis, 2017; Vogel, 2016).

Despite the lack of research into this particular field, the analysis that does exist on the internet’s role in politics provides context for my research. Before the internet was studied as a

vessel for fundraising, it was seen more broadly as an agora of public discussion (di Gennaro & Dutton, 2006; Gallego & Cantijoch, 2010; Gil de Zúñiga, Puig-I-Abril, & Rojas, 2009; Valenzuela, Kim, & Gil de Zúñiga, 2011; Vesnic-Alujevic, 2012), a hub of political organizing (Jost, et al., 2018; Cogburn & Espinoza-Vasquez, 2011; Levenshus, 2010), and predictor of offline political capital (Gil de Zúñiga, Jung, & Valenzuela, 2012; Hardina, 2005). The bulk of literature that explicitly focuses on online fundraising comes from the field of nonprofit organizational studies (Hazard, 2003; Karpf, 2010 Marx DSW, 2008; Miller, 2009; Raihani & Smith, 2015).

Although the study of the internet as a medium is relatively new to political science, research suggests that its communication methods are similar to traditional political communication and can be extrapolated to offline characteristics. The differences that are seen in online political communication, like lowered costs and eased barriers to entry, represent a “difference-of-degree” and not a paradigm shifting “difference-in-kind” (Karpf, 2010). Furthermore, there is a strong connection between online channels of communication in the form of social networks and offline connections and building and maintaining social capital from those offline connections (Cranshaw, Toch, Hong, Kittura, & Sadeh, 2010; Ellison, Steinfield, & Lampe, 2006; Liben-Nowell, Novak, Kumar, Raghavan, & Tomkins, 2005; Scellato, Mascolo, Musolesi, & Latora, 2010). And online social networks have been used to study offline-based actions and beliefs like opinion polarization (Lee, Choi, Kim, & Kim, 2014), political polarization (Hanna, et al., 2013), political participation (Lawrence, Sides, & Farrell, 2010) and political discourse (Kushin & Kitchener, 2009). The bottom line is that online actions and behaviors have been found to mirror the offline world and they are frequently extrapolated to explain offline actions and behaviors by other studies in political science and communication. My research builds upon these previous

uses of online indicators of offline actions and beliefs to layer in political elites' online communications to decipher the offline ideologies of non-elite actors.

Given the inherent connection between online and offline actions, the more traditional study of the role of geography in politics can be insightful in studying both online networks and the motivations of political donors. It is well documented that the American electorate self-sorts geographically. Whether this self-sorting is a central factor in geographic relocation or an ancillary byproduct, liberals tend to move to geographic areas where there are more liberals, and conservatives live near other conservatives (Cho, Gimpel, & Hui, 2013). This partisan sorting leads to small-group conformity in which individuals receive personal social rewards by adhering or otherwise conforming to group norms (Bernheim, 1994). This social phenomenon of conformation exists in political donations. One study found that when an individual's political contributions are made more visible to their neighbors, contributions to the local majority party increase and contributions from supporters of the minority party decrease (Perez-Truglia & Cruces, 2017). In other words, it is possible that geographic clustering of donors is a self-fulfilling prophecy because local political conditions impact one's donation decisions. But it is not clear whether individuals act to avoid social sanctions or gain social standing despite their personal beliefs, or if a person's beliefs (and subsequent political action) is shaped by their local social networks (Perez-Truglia & Cruces, 2017). In other words, scholars do not agree upon why people self-sort geographically. But it is a consensus view that partisans do sort themselves.

This self-sorting creates geographic networks of political contributions that are wide-spread and independent of wealth, age, occupation and other individual characteristics. Contributors are "part of networks in which social influence can be brought to bear in the solicitation of contributions" (Gimpel, Lee, & Kaminski, 2006). In addition, "Republican and Democratic donor bases are much more similar geographically than their bases of electoral support." Spatial

proximity creates a greater sense of perceived “common material interest” in government policy. And therefore, individuals are more likely to organize (or make political contributions) to achieve that perceived local interest (Gimpel, Lee, & Kaminski, 2006). As it relates to my research into the intersection of online political communication and geography, it is put best that there is an “exaggerated death of geography,” and that we should question “the ‘distance-destroying’ capacity of information and communication technologies where social depth is conflated with spatial reach” (Gimpel, Lee, & Kaminski, 2006). Essentially, “physical proximity may be essential for some forms of knowledge exchange” because in order to fully commit to a cause or build trust, there must be a “prolonged socialization and regular reassurance of the other person’s sincerity” (Gimpel, Lee, & Kaminski, 2006; Morgan, 2004). Political donors should not be thought of as simplistic actors who act in a vacuum. They are generally a part of a larger pool of potential resources that are geographically conditioned by their social networks to provide and seek out bonds of trust on issues of common material interest before making a contribution. These bonds of trust are likely to be best communicated by candidates at a local level (Gimpel, Lee, & Kaminski, 2006).

Local candidates can talk about issues of common material interest, otherwise thought of as solidarity—“people often contribute out of a desire to feel that they are a member of a high-status social group or part of a political team fighting for a cause” (Gimpel, Lee, & Kaminski, 2006). And so, it not surprising that there is a strong connection between local political networks and solicitation of political contributions (Brady, Schlozman, & Verba, 1999). The power of these personal networks can be thought of as a function of the “strength, immediacy, and number of persons standing together” (Latané, 1981). The recent emergence of online donor communities could theoretically leave much of this past research moot. The study of geography in political elections was deeply tied to party officials and local elites’ ability to capitalize on

their social standing through their immediate geographic social networks. These local elites were able to utilize their stature and immediacy because donors were believed to greatly value personal appeals that are done in face-to-face (Alexander, 1992; Brown Jr., Powell, & Wilcox, 1995). As a result, primarily online donors would likely have a motivation to contribute other than in-person interactions and the networks that it formed.

On the other hand, the previously-mentioned research that donors are more motivated by solidarity to local issues could take on greater significance. My study looks at state legislative candidates who in theory could personalize their communications to speak to local issues that could emphasize common material interest. And social media could allow them to reach a more niche audience to form their own messaging to potential donors instead of a common partisan message from systems of mass communication.

Although a minority, an alternative view exists that the consideration of geography in politics forces arbitrary boundaries to be drawn. And therefore, “context should not count,” and that one’s geographic socialization does not significantly change one’s political actions. The main argument is that, “Political scientists need political geographers because they are skillful at pointing out what we do not understand. Geographic tools are essential for displaying areal variation in what we know, but this is nowhere near as powerful as the role of geography in revealing features of data and the political world that we would not otherwise have considered” (King, 1996). A recent event that could solidify this argument is the 2017 special election for Georgia’s sixth congressional district. The race became the most expensive one in U.S. history because of the \$28 million raised by the candidates. The race took on a nationalized agenda that saw a large number of donations coming from outside of Georgia. Only 11 percent of itemized contributions to the Democrat Jon Ossoff were from the state. The race was reported to have become nationalized (Parlapiano & Shorey, 2017). And it seems unlikely that the significant

number of non-Georgia donors had a sincere connection to the local political climate and common material interests that shaped local political issues.

Previous studies into these three different theories of donor motivation (motivated by access, motivated by geographic network, motivated by something other than geographic network, but not necessarily seeking access) treats each theory as if they are in a binary vacuum—either the single explanation they are testing is true or false. Instead, my research employs a methodology that can test all three at the same time. Furthermore, my research recognizes that different communities of donors could have separate interests and is theoretically able to identify each community’s motivations. Perhaps some communities of donors are access-oriented, maybe other are more geographically based, and some could be motivated by certain policies. My research questions are: Do different groups of donors have different motivations? If some groups of donors seem to be motivated by candidate communications, what are those issues?

Measuring Ideology

Fundamentally, my research is a methodology paper that proposes an experimental and novel way to consider the intersections between political donor communities, social media communication and geography. The inspiration for attempting such research draws from previous attempts to uncover ideology and policy preferences of political actors using quantitatively-driven methods.

As William Berry, one of the leading scholars in ideology estimation, put it in one of his papers, “Few concepts are more central to politics and policy than the ideology of citizens...but state politics and policy scholars have been particularly interested in citizen ideology as an agent of policy change” (Berry, Ringquist, Fording, & Hanson, 2007). One could see the key to understanding changes in public policy as recognizing shifts in citizens’ ideologies.

The debate on the best method to estimate the ideology of the electorate has traditionally broken down into two camps: those who use observational data to make inferences and those who primarily rely upon surveys. Each side has its proponents and detractors. Previous studies found compelling results in the estimation of U.S. states' ideologies by looking at the outcomes of congressional elections, roll call votes of congressional delegations, interest-group ratings for members of Congress, and the partisanship and results of state elections (Berry, Ringquist, Fording, & Hanson, 1998; Holbrook-Provow & Poe, 1987). Meanwhile, survey-based research claimed to be able to make similar ideological estimations of aggregate populations by extrapolating polling data (Brace, Arceneaux, Johnson, & Ulbig, 2004; Wright, Erikson, & McIver, 1985). Eventually, Berry et al. concluded that inferred ideology and survey responses measured two fundamentally different things. They argued that results from action-observed methodologies indicate political mood, whereas surveys are a measurement of self-identified ideology. Unsurprisingly, they maintained that their non-survey method "is more appropriate when studying the impact of public opinion on public policy" because it's based on people's actual actions and not just their reported (but possibly unacted upon) feelings (Berry, Ringquist, Fording, & Hanson, 2007).

Beyond voters' ideology, the ideology of political donors is also an ongoing topic in political science research. Hill and Huber point out that "we know relatively little about the contemporary representativeness of those who donate compared to the larger American electorate" despite money in American politics impacting who runs for office, helping to determine which candidates win elections, and the public fears about the influence of money on the political process (Hill & Huber, 2016).

The most comprehensive study of political donors' ideology is Adam Bonica's "Mapping the Ideological Marketplace" (Bonica, 2014). Bonica used over 100 million contribution records

from state and federal elections to create ideological ideal point estimates for every candidate and donor in his data. The comprehensiveness of his data allows him to map virtually any person that has made a political contribution between 1972 and 2012 onto a “common space” with candidates and PACs to make coherent liberal-conservative ideological estimations across a wide variety of political actors.

While not the focus of his paper, he does incorporate donors’ policy preferences as a part of his discussion in his findings. He uses the example of the 2012 Republican presidential primary candidates and examines how pro-choice or pro-life their donors were. He estimates their pro-choice/ pro-life level by classifying each contributor based on whether they donated to a pro-life organization or ballot initiative (e.g., Right to Life) or a pro-choice organization or ballot initiative (e.g., NARAL). He then calculated each donors’ “pro-life proportions” by dividing their total pro-life contributions by the total of their pro-life and pro-choice contributions. This simplicity works for his ad hoc analysis of pro-life donors in the 2012 Republican presidential primary, but it has some holes when using it to make broad judgements. For example, it does not consider donors who contribute to candidates. Although candidates campaign on many policy issues, it cannot be discounted that donors contribute to campaigns solely based on one issue. In other words, donors could have contributed to candidates solely because they supported that candidate’s policy on being pro-life or pro-choice, but that would not have been picked up in Bonica’s explanation. My research seeks to fill this research gap of identifying the issues that motivate political donors by using a similar “common space” approach as Bonica but adds network analysis and social media data to determine the issues that political candidates publicly champion to attempt to uncover the policy issues that motivate political donors.

However, Bonica’s methodology has been challenged. Hill and Huber compared Bonica’s ideology scores to survey data from the 2012 Cooperative Congressional Election

Survey (CCES). They concluded that only examining political donations did not do an adequate job of estimating ideology on a spectrum. They conceded that donations could be helpful for sorting people into Democratic and Republican groups but did a poor job of placing donors on an ideological spectrum intraparty. They concluded that this is the case because candidate and donor ideology are only weakly connected, and that even low-dollar donors contribute to competitive campaigns where their donations can have the biggest impact (Hill & Huber, 2016).

Bonica responded by validating his own scores based on direct comparisons of 30 CCES policy items. He stated that, ultimately, his ideology scores are “powerful predictors of policy preferences for a wide range of issues and successfully discriminate within party” (Bonica, 2017).

While my research is most closely related to Bonica’s, it is persuaded by Hill and Huber’s findings in that it utilizes a community-based approach that does not rely on exact individual placement along a continuum. Placement in my spatial network is only significant in determination of statistical communities. Where Bonica uses this placement to extract ideology on a left-right continuum, I layer on social media data to create policy ideologies of the aggregate communities. Social media data has not widely been used in this fashion. But my findings show that it can be used as an additional variable in the understanding of networks of social connections between elite and non-elite political actors.

Method

To determine if there is a correlation between communities of donors and the policy issues in the online communications of political candidates, I used an original and novel methodology. I took four principle steps: creation of donor community scores, creation of online communication topic scores, traditional statistical analysis determining if there is a correlation between the donor

scores and topic scores, and geographic cluster analysis to determine if geography is an explanation for any of the donor communities.

Donor Scores

To create donors scores, I started with all 92,807 political contributions to candidates for Wisconsin State Senate and State Assembly in 2015 and up to the fall 2016 election. These donations were pulled from the Wisconsin Campaign Finance Information System (Wisconsin Ethics Commission, n.d.). I then removed any contributions made anonymously (currently, contributions under \$50 can be given anonymously) or unitemized. This step left 92,284 donations.

Next, I made the data uniform (removed punctuation, made all names lowercase, etc.) and used OpenRefine (Ham, 2013) to stem names to identify people who might be the same person, but were entered differently. For example, if one campaign reported a contribution from Jim Smith, and another reported a donation from James Smith, both records were changed to be identical (e.g., Jim Smith). To ensure that people who had the same name, but were different people, were not counted as the same individual, I added their zip code to the end of their name to create a unique identifier (e.g., Jim Smith: 01234). I manually checked to see if it would be an issue that multiple people in the same zip code would get stemmed to the same name and found that this was not a concern for the analysis. I then tallied up how many contributions came from each unique identifier. I kept only the donations from people that made more than one contribution. I did this for two main reasons. First, removing these excessive donors allowed for more computational efficiency. Second, I wanted to focus on donors with potentially identifiable policy preferences. It would be difficult to identify the policy issues that motivates donors that gave a single donation because there are so many different reasons that someone could make a

donation that has nothing to do with policy. For example, I could donate money to a family member or friend who is running for office. Instances like this one would be much less likely in people who gave multiple contributions and therefore are more engaged in the political donation process. It should be noted, that I kept donations from people who gave multiple times, but only to one candidate. This step left me with 29,990 donations.

After this data preparation process, I calculated the statistical communities within the donor network using Gephi (Bastian, Heymann, & Mathieu, 2009). In essence, the political donor landscape can be thought of as a network of nodes (donors and politicians) who are connected by edges (political contributions). We can then use modular community detection to identify smaller statistical clusters within the community based on the number of shared connections between nodes (see Figure 1). Specifically, I utilized Newman's modularity—the most widely used modularity detection method (McSweeney, 2009). For example, if nodes A, B, C, and D are all connected to each other, but only share sparse connections to the rest of the network, they are identified as being in a distinct community. To best visualize the network, I used the Yifan Hu layout algorithm (Hu, 2005). This algorithm is force-directed—meaning that it tries to polarize the network away from each other, but nodes are kept more compact when they share more connections. In this application, we see that the network forms two polarized clusters: Democrats and Republicans. For the sake of this analysis, the most important component is noting the statistical cluster that every donor belongs to because my hypothesis rests on these clusters sharing some identifiable common characteristic like donating to candidates who talk about similar policy issues or contributing to candidates in a geographic area (see Figure 1).

I used these clusters to create donor cluster scores for every candidate. I kept the calculation of these scores simple. However, more sophisticated modeling could be used in the future or if a similar analysis is done at a larger scale. For every candidate, I calculated the percent of their

donations that came from each cluster. I then multiplied that number by the percent that that cluster's donations would get the candidate to a "competitive fundraising total." To determine the total amount that a campaign would need to raise to get to a "competitive fundraising total," I took the median contribution total to campaigns in competitive seats. I used Bonneau and Hall's definition that elections in which no candidate won more than 60% of the vote as being competitive (Bonneau & Hall, 2003). In this case, I took the medians of 2016 fundraising totals for seats that were won with less than 60% of the vote in 2014. For State Senate elections that were competitive, the median competitive campaign raised \$336,868.93. Competitive State Assembly campaigns had a median of \$57,446.74 in contributions. The formula looks like this: $((\text{amount raised from donor cluster} / \text{total amount raised} \times 100) \times (\text{amount raised from donor cluster} / \text{competitive fundraising total} \times 100))$. For example, say that Jane Doe for Assembly raised \$20,000 in total, \$10,000 of which came from donors in donor cluster 1. Jane Doe's donor cluster 1 score would be 5.5 $((\$10,000 / 20,000 \times 100) \times (\$10,000 / 54,449.20 \times 100))$ I calculated these donor scores in this fashion to balance the potential impact of a single donor community on the campaigns while also considering the scale of the donations. If a single community was responsible for 100% of a candidate's donations, that candidate is strongly connected to that community. However, if that candidate only raised \$1,000, that community's impact should be treated less than a candidate who raised \$200,000 and received 70% of their contributions from a single community.

Topic Scores

After I calculated donor scores for every candidate, I moved my focus towards creating social media communication topic scores for every candidate based on the issues that they posted about on Twitter and Facebook. I chose to analyze social media data because it the most accessible and quantifiable data that is available on campaigns' communication efforts. In

addition, I hypothesized that social media provides the best route for candidates to attract unique communities of donors regardless of geography.

First, I manually collected the Twitter handles and Facebook usernames of every candidate running for Wisconsin State Assembly and State Senate. I included both official legislative social media accounts and campaign accounts. Although this might skew the number of posts made by politicians in office (who have both official legislative and campaign accounts), this holistic approach best captures the entirety of a political donors' potential experience with a candidates' online presence. I used this list of accounts to collect all tweets and Facebook posts from these accounts from January 1, 2015 through November 8, 2016 (election day) through the Twitter (Kearney, 2017) and Facebook (Barbera, Piccirilli, Geisler, & van Atteveldt, 2017) APIs. In total, I collected 82,851 posts.

To determine the issues that each candidate talked about on social media, I trained a neural network in R using the *quanteda* package (Benoit, 2018). To create a training and test set, I hand-coded a randomized 15%—12,428—of the posts (10% for a training set, 5% for test set). I inductively created a codebook based on the topics that were being posted about—the same method employed by some of the most famous political science studies (Hershey, 1992). I made sure to differentiate between the view point that the post represented. For example, there was not just one “marijuana” category, instead, there were “pro-marijuana” and “anti-marijuana” categories. After coding all of the posts, I collapsed categories with fewer than 20 posts, or .16% of total posts, into the “other” category. In a perfect world, I could have kept these categories. However, there just weren't enough training posts for the neural network to properly be able to categorize posts in these categories. For the sake of accuracy in the model, I had to put them into the “other” category. I also collapsed categories that fell under similar themes but differed slightly in specific policy. For example, I originally had, “pro-right-to-work” and “anti-union”

categories that were combined into a more general “conservative on economic issues” category. I ultimately ended up with 28 categories.

I implemented a supervised latent dirichlet allocation model to train a neural network to classify the posts (Blei & McAuliffe, 2008). As with the selection of any machine learning algorithm, there are pros and cons. Latent dirichlet allocation models are generally used on data sets with fewer topics than my 28. However, it performs best when there is a small number of specific words that make the text document likely to be in a topic—a methodology that works well with short texts like tweets and Facebook posts.

I used my training set to train the model to classify posts based on the classifications I gave the posts in my training set. I then measured the accuracy of my model by running it on my pre-coded test set and comparing whether the model’s output matched my hand coding. The model was not extremely accurate but performed admirably with 57.58% accuracy. This number may seem low. However, it is considerably better than if the model just assigned outputs at random (that would have a 3.57% accuracy). In addition, the model seemed to get close, but not quite there on many topics. For example, I classified this post, “Today @JonErpenbach & I unveiled bill that offers veterans immediate mental health care svcs. during VA wait times” as being in the “veterans issues/ King veterans home” category, but the model classified it as “liberal healthcare”—a discrepancy that could be argued either way if done by hand coders. I was concerned that the model’s minor discrepancies on individual posts would adversely impact the next step of creating topic scores for each candidate. However, the scores held face validity. For example, Representative David Bowen (D-Milwaukee), a well-known civil-rights activist had a high topic score for “race issues,” Representative Melissa Sargent (D-Madison), an advocate of women’s health scored a high topic score for “pro-choice/ women’s health,” and Representative John Nygren (R-Marinette), the main proponent of the Heroin, Opioid Prevention and Education

(HOPE) Agenda, and whose daughter has publicly struggled with substance abuse, received a high topic score for “drug abuse and prevention.”

While I admit that the model is not perfect, machine learning—especially applied to natural language processing—remains an extremely experimental field, and this method is the best that I could implement. I attempted to closely follow other ground-breaking work in this area (Yano & Smith, 2010) and could be a beneficiary of other ongoing research into text classification from other researchers (Quan, Zhang, Xu, & Hossain, 2015). Hopefully, in the future, a more accurate and easily implementable method for classification will be developed. But for now, my model is accurate enough when taken in aggregate to continue my analysis.

I calculated topic scores in a nearly identical manner to the donor scores. I took the percent of a candidate’s posts that were classified as belonging to each topic and multiplied it by the number of posts on each topic divided by the “competitive post total” (the median number of posts made the competitive campaigns) of 210. The formula is: $((\text{posts about topic A} / \text{total posts} \times 100) \times (\text{posts about topic A} / \text{competitive post total} \times 100))$. For example, if Jane Doe made 100 posts about topic A out of 200 total posts, her topic A score would be 7.6 $((100 / 200 \times 100) \times (100 / 148.5 \times 100))$. Again, this simple method was employed to balance how much a campaign focuses on a single topic versus scale. If a single topic makes up only a small percent of a candidate’s posts, but that candidate posts so much that the raw number is greater than other campaign’s total posts, that candidate’s topic score should be reflected as being high.

Statistical Analysis

Next, I statistically analyzed to see if there was a correlation between any of the donor and topic scores. First, I calculated a correlation coefficient for every combination of donor and topic scores (see Figure 3). Next, I calculated the statistical significance of the combinations by first

calculating their p-values and converted those to an adjusted p-value. I opted to convert the p-values to adjusted p-values because I wanted to control for a false discovery rate since I was computing a large number of—364—p-values (see Figure 4). I then overlaid the statistically significant donor groups and topic combinations (those with a p-value < .05) (see Figure 5).

Geographic Clustering

Last, I hypothesized that some of the clusters' common identifiable trait is that they are geographically clustered. To measure geographic clustering, I converted every donors' address to latitude and longitude and placed them on a map. I then calculated the standard distance for each cluster's donors. The location of every individual donor in each cluster can be seen in Figure 6 and the standard distance of each cluster in Table 1. In essence, standard distance creates a measurement of compactness of points. The lower the standard distance, the more compact the points are, the higher the standard distance, the more dispersed (ArcGIS, 2017).

Results

My analysis set out to uncover the issues that motivate communities of political donors who contribute to campaigns at a state level. Because of the way that the internet has transformed political fundraising, I expected that candidates' ability to engage individuals on the internet—regardless of geography—would mean that donors would coalesce into communities by contributing to candidates who are champions of specific policies.

Taken by itself, I have found statistically compelling evidence that there is a correlation between the amount of money raised by campaigns from various donor groups and specific policy issues that those candidates discuss on social media. In addition, there is a clear role that geography plays in the creation of some of the statistically identifiable donor communities.

Therefore, stating that social media communications from campaigns is the only motivating factor behind donor contributions would be an oversimplification.

Although the relationship between donors and policy issues cannot be directly stated, I have found that geographic considerations should play a larger role when studying the issues that campaigns decide to engage with. Even in the age of the internet and mass communication, we should not discount the old axiom that ‘all politics is local’ because every locality seems to have distinct issues that their political leaders discuss and potentially motivate citizens to engage with the civic process—specifically make a political contribution.

When considering the geography (Figure 6), correlated topic scores (Figure 5), and competitiveness of the races the donors contributed to (Figure 7), three main types of donor clusters emerge: groups with only one or two policy motivations; groups of donors that represent “base” donors; and donors with geographic clustering but no correlation with policy topics.

The first category of donor clusters is those whose contributions are statistically-significantly correlated with candidates talking about one or two specific policy issues. One of the strongest examples of this type of behavior is Cluster 2. These donors are geographically spread out across Wisconsin, contributed primarily to Republicans, and contributed about half of their money to competitive campaigns. Geography, nor competitiveness, seem to be connecting explanations of these donors’ behavior. However, there is a strong correlation between campaigns receiving more contributions from this group of donors and campaigns talking about bipartisan issues in healthcare (R^2 of $\sim .7$). It is possible that these donors are Republican donors who seek out candidates who are moderate or pragmatic on healthcare.

An example of a Democratic-donating cluster is Cluster 3. These donors are also spread-out across Wisconsin and contributed about two-thirds of their money to competitive races.

However, they have a strong correlation to the environment and climate change topic scores.

Another unique example is donor Cluster 8. These donors are from all over Wisconsin, contributed about half of their money to competitive campaigns, and their donations were correlated to candidates talking about drug abuse and anti-gun control. This donor cluster was the only cluster with a statistically significant correlation to either of these two policy issues.

Two other clusters fall into this single-issue classification contribute to Democrats but have possible non-policy explanations for their donations. Cluster 1 has a statistically significant relationship with criminal justice reform. However, these donors are clustered tightly in western South Central Wisconsin—their standard distance (Table 1) is orders of magnitude the lowest compared to the other clusters—and they contributed 98.4% of their money to competitive races. These factors suggest there was more than just a single policy issue that connects these donors. But these considerations allow for some compelling theories to be developed. Possibly, criminal justice reform was a popular topic among Democrats in this specific geography. Or, these competitive campaigns saw the issue as a winning one in the general election. Another cluster, Cluster 11, is unique in that it is only correlated to pro-LGBTQ+ issues. However, their donors are primarily confined within the borders of Milwaukee County. Cluster 11, along with Cluster 6 (Racine and Kenosha Counties), are the only clusters to show such a profound geographic pattern with respect to political borders. Again, Cluster 11 elicits more questions. Are donors from Milwaukee County particularly pro-LGBTQ+, even compared to other strongly Democratic areas? Are candidates from Milwaukee themselves that much more publicly supportive of the issue in their social media communications than candidates from other parts of the state such as Madison? Investigations into these two donor clusters alone could be the topic of future research.

The next category of donors can be best described as base donors. They are all from predominately strong geographic areas of support (Republican Cluster 0 from Southeastern

Wisconsin, including the Milwaukee County suburbs; Democratic Cluster 6 from Racine and Kenosha Counties; and Democratic Cluster 10 from Madison/ Dane County). These donors from traditionally partisan areas did not donate very much money to competitive campaigns, instead contributing primarily to their local non-competitive races (21.2%, 11.9%, and 25.6% of each cluster's money was contributed to competitive campaigns, respectively). And their donations were significantly correlated with several topic issues. Donations from Cluster 0 were correlated with candidates talking about pro-life issues, conservative education, conservative economic issues, and criminal justice/ police. Cluster 6 was correlated with liberal voting (access to voting), the environment and climate change, liberal education, liberal economic issues, and campaign finance reform/ corruption/ open government. Cluster 10 was strongly correlated with pro-choice/ women's health along with significant correlations to liberal voting, pro-marijuana, pro-LGBTQ+, veterans issues, liberal healthcare, pro-gun control, anti-gerrymandering, liberal education, liberal economic issues, and campaign finance reform/ corruption/ open government. Even though all three of these clusters represent a variety of policy topics, their differences can be informative. For example, comparing the Madison donor cluster (Madison has a reputation as being very liberal, even compared to other Democratic areas) to the Racine and Kenosha Counties group, donations from the Madison group are correlated with issues that might be seen as more "socially liberal" such as being pro-marijuana, pro-choice/ women's health, and pro-LGBTQ+. These issues are not correlated with the other "base" donor group—showing the intra-party distinction of policy issues that is unique by geography.

The final category of donor clusters is those that are not correlated with any specific policy issues but are clustered by geography. Cluster 4 is made up of Democratic donors primarily from Central Wisconsin, including strongly Democratic college town of Stevens Point. Cluster 5 is a Republican cluster from Southwestern Wisconsin. Cluster 7 is comprised of Democratic donors

from Northwestern Wisconsin. Cluster 9 is a Republican cluster from Northwestern Wisconsin. And Cluster 12 is made up of Democratic donors from Southwestern Wisconsin.

In many ways, this is a methodological paper that set out to experiment with a novel way of identifying the intersection between political donor communities, geography and social media communications to construct donor issue publics. And a large amount of output data was created. I invite readers to use this data to come up with their own hypotheses and ask their own questions such as digging into specific policy topics or geographies of interest.

Discussion

The role the internet played in the 2016 presidential election was the impetus for this research and its attempt to use social media data to infer the issues that political donors care about. This research was able to identify statistically significant correlations between donor clusters and candidates talking about specific topics on social media. Another significant finding is that geography and the issues that impact physical communities should play a larger role in the practice and study of political communication.

Both scholars (Cramer, 2016) and popular authors (Vance, 2016) have brought the importance of geography to the forefront of political communication research. But these discussions rely on a rural-urban dichotomy that my research shows may not even be granular enough. For example, rural communities in southwestern Wisconsin may have priorities that differ from ruralites in northeastern Wisconsin that are greater than their general “dislike” for urban centers. Further, this research shows geography and localism are not only important in traditional political science but is also manifested in political communication—specifically in digital communication.

Practically, this research could inform the strategy of political campaigns. Both local candidates could quickly understand the issues that their campaigns may wish to prioritize, but state-wide or even national campaigns could better understand the policy issues to stress when visiting or advertising in the region.

Overall, this paper contributes to the burgeoning field of ideological estimation. It provides novel methodology that successfully finds a correlation between specific donors contributing to campaigns and campaigns discussing various policy issues. It also highlights the significance that geography plays in traditional political science and the influence it can have on political communication.

References

- Alexander, H. E. (1992). *Financing Politics: Money, Elections, and Political Reform*. Washington: CQ Press.
- Ansolabehere, S., de Figueiredo, J. M., & Snyder Jr., J. M. (2003). Why is There so Little Money in U.S. Politics? *Journal of Economic Perspectives*, 105-130.
- ArcGIS. (2017). How Standard Distance works. Retrieved from ArcGIS Pro: <http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-standard-distance-spatial-statistic-works.htm>
- Bachrach, P., & Baratz, M. S. (1962). Two Faces of Power. *The American Political Science Review*, 947-952.
- Barber, M. (2016). Donation Motivations: Testing Theories of Access and Ideology. *Political Research Quarterly*, 148-159.
- Barbera, P., Piccirilli, M., Geisler, A., & van Atteveldt, W. (2017). Rfacebook: Access to Facebook API via R version 0.6.15. Retrieved from <https://cran.r-project.org/web/packages/Rfacebook/Rfacebook.pdf>
- Bastian, M., Heymann, S., & Mathieu, J. (2009). Gephi: an open source software for exploring and manipulating networks. *Icwsn* 8, 361-362.
- Benoit, K. (2018). quanteda: Quantitative Analysis of Textual Data version 1.0.0. Retrieved from <https://cran.r-project.org/web/packages/quanteda/quanteda.pdf>
- Bernheim, B. D. (1994). A Theory of Conformity. *Journal of Political Economy*, 841-877.

- Berry, W. D., Ringquist, E. J., Fording, R. C., & Hanson, R. L. (1998). Measuring Citizen and Government Ideology in the American States. *American Journal of Political Science*, 327-348.
- Berry, W. D., Ringquist, E. J., Fording, R. C., & Hanson, R. L. (2007). The Measurement and Stability of State Citizen Ideology. *State Politics and Policy Quarterly*, 111-132.
- Blei, D. M., & McAuliffe, J. D. (2008). Supervised Topic Models. *Advances in Neural Information Processing Systems*, 121-128.
- Bonica, A. (2014). Mapping the Ideological Marketplace. *American Journal of Political Science*, 367-386.
- Bonica, A. (2017, December 12). Are Donation-Based Measures of Ideology Valid Predictors of Individual-Level Policy Preferences. Retrieved from stanford.edu: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3087782
- Bonneau, C. W., & Hall, M. G. (2003). Predicting Challengers in State Supreme Court Elections: Contet and the Politics of Institutional Design. *Political Research Quarterly*, 337-349.
- Brace, P., Arceneaux, K., Johnson, M., & Ulbig, S. G. (2004). Does State Political Ideology Change Over Time. *Political Research Quarterly*, 529-540.
- Brady, H. E., Schlozman, K. L., & Verba, S. (1999). Prospecting for Participants: Rational Expectations and the Recruitment of Political Activists. *American Political Science Review*, 153-168.
- Briffault, R. (2010). Corporations, Corruption, and Complexity: Campaign Finance After Citizens United. *Cornell Journal of Law and Public Policy*, 643-672.
- Brown Jr., C. W., Powell, L. W., & Wilcox, C. (1995). *Serious Money: Fundraising and Contributing in Presidential Nomination Campaigns*. Cambridge: Cambridge University Press.
- Brunell, T. L. (2005). The Relationship Between Political Parties and Interest Groups: Explaining Patterns of PAC Contributions to Candidates for Congress. *Political Research Quarterly*, 681-688.
- Cho, W. K., Gimpel, J. G., & Hui, I. S. (2013). Voters Migration and the Geographic Sorting of the American Electorate. *Annals of the Association of American Geographers*, 856-870.
- Coates IV, J. C. (2012). Corporate Politics, Governance, and Value Before and After Citizens United. *Journal of Empirical Legal Studies*, 657-696.
- Cogburn, D. L., & Espinoza-Vasquez, F. K. (2011). From Networked Nominee to Networked Nation: Examining the Impact of Web 2.0 and Social Media on Political Participation and Civic Engagement in the 2008 Obama Campaign. *Journal of Political Marketing*, 189-213.
- Cramer, K. J. (2016). *The Politics of Resentment: Rural Consciousness in Wisconsin and the Rise of Scott Walker*. Chicago: University of Chicago Press.

- Cranshaw, J., Toch, E., Hong, J., Kittura, A., & Sadeh, N. (2010). Bridging the gap between physical location and online social networks. *ACM international conference on Ubiquitous computing* (pp. 119-128). New York: ACM.
- di Gennaro, C., & Dutton, W. (2006). The Internet and the Public: Online and Offline Political Participation in the United Kingdom. *Parliamentary Affairs*, 299-313.
- Ellison, N., Steinfield, C., & Lampe, C. (2006). Spatially Bounded Online Social Networks and Social Capital: The Role of Facebook. *Annual Conference of the International Communication Association*. Dresden, Germany: International Communication Association.
- Epstein, R. A. (2011). *Citizens United v. FEC: The Constitutional Right That Big Corporations Should Have But Do Not Want*. *Harvard Journal of Law & Public Policy*, 639-662.
- Foran, C. (2016, March 1). Bernie Sanders's Big Money. Retrieved from The Atlantic: <https://www.theatlantic.com/politics/archive/2016/03/bernie-sanders-fundraising/471648/>
- Gallego, A., & Cantijoch, M. (2010). Online Political Participation in Spain: The Impact of Traditional and Internet Resources. *Journal of Information Technology & Politics*, 356-368.
- Gerken, H. K. (2014). The Real Problem with Citizens United: Campaign Finance, Dark Money, and Shadow Parties. *Marquette Law Review*, 903-924.
- Gil de Zúñiga, H., Jung, N., & Valenzuela, S. (2012). Social Media Use for News and Individuals' Social Capital, Civic Engagement and Political Participation. *Journal of Computer-Mediated Communication*, 319-336.
- Gil de Zúñiga, H., Puig-I-Abril, E., & Rojas, H. (2009). Weblogs, traditional sources online and political participation: an assessment of how the internet is changing the political environment. *New Media & Society*, 553-574.
- Gimpel, J. G., Lee, F. E., & Kaminski, J. (2006). The Political Geography of Campaign Contributions in American Politics. *The Journal of Politics*, 626-639.
- Gordon, S. C., Hafer, C., & Landa, D. (2007). Consumption of Investment? On Motivations for Political Giving. *The Journal of Politics*, 1057-1072.
- Hall, A. B. (2014, January 13). How The Public Funding of Elections Increases Candidate Polarization. Retrieved from [andrewbenjaminhall.com](http://www.andrewbenjaminhall.com): <http://www.ifs.org/wp-content/uploads/2014/07/Hall-2014-Tax-Financing-And-Polarization.pdf>
- Ham, K. (2013). Open Refine (version 2.5). <http://openrefine.org>. Free, open-source tool for cleaning and transforming data. *Journal of the Medical Library Association: JMLA* 101.3, 233.
- Hanna, A., Wells, C., Maurer, P., Friedland, L., Shah, D., & Matthes, J. (2013). Partisan alignments and political polarization online: a computational approach to understanding the french and US presidential elections. 2nd workshop on Politics, elections and data (pp. 15-22). San Francisco: PLEAD.
- Hardina, D. (2005). Using the Web to Teach Power Analysis. *The Social Policy Journal*, 51-68.

- Hazard, B. L. (2003). Online fundraising at ARL libraries. *The Journal of Academic Librarianship*, 8-15.
- Hershey, M. R. (1992). The Constructed Explanation: Interpreting Election Results in the 1984 Presidential Race. *The Journal of Politics*, 943-976.
- Hill, S. J., & Huber, G. A. (2016). Representativeness and Motivations of the Contemporary Donorate: Results from Merged Survey and Administrative Records. *Political Behavior*, 3-29.
- Holbrook-Provow, T. M., & Poe, S. C. (1987). Measuring State Political Ideology. *American Politics Quarterly*, 399-416.
- Hu, Y. (2005). Efficient, high-quality force-directed graph drawing. *Mathematica Journal* 10.1, 37-71.
- Jost, J. T., Barberá, P., Bonneau, R., Langer, M., Metzger, M., Nagler, J., . . . Tucker, J. A. (2018). How Social Media Facilitates Political Protest: Information, Motivation, and Social Networks. *Political Psychology*, 85-118.
- Kang, M. S. (2011). After Citizens United. *Indiana Law Review*, 243-254.
- Karpf, D. (2010). Online Political Mobilization from the Advocacy Group's Perspective: Looking Beyond Clicktivism. *Policy & Internet*, 7-41.
- Kearney, M. W. (2017). rtweet: Collecting Twitter Data. R package version 0.6.0. Retrieved from <https://cran.r-project.org/web/packages/rtweet/citation.html>
- King, G. (1996). Why context should not count. *Political Geography*, 159-164.
- Krueger, B. S. (2002). Assessing the Potential of Internet Political Participation in the United States. *American Politics Research*, 476-498.
- Kushin, M. J., & Kitchener, K. (2009). Getting political on social network sites: Exploring online political discourse on Facebook. *First Monday*.
- La Raja, R. J., & Schaffner, B. F. (2014). The effects of campaign finance spending bans on electoral outcomes: Evidence from the states about the potential impact of Citizens United v. FEC. *Electoral Studies*, 102-114.
- Langbein, L. I. (1986). Money and Access: Some Empirical Evidence. *The Journal of Politics*, 1052-1062.
- Latané, B. (1981). The psychology of social impact. *American Psychologist*, 343-356.
- Lawrence, E., Sides, J., & Farrell, H. (2010). Self-Segregation or Deliberation? Blog Readership, Participation, and Polarization in American Politics. *Perspectives on Politics*, 141-157.
- Lee, J. K., Choi, J., Kim, C., & Kim, Y. (2014). Social Media, Network Heterogeneity, and Opinion Polarization. *Journal of Communication*, 702-722.
- Levenshus, A. (2010). Online Relationship Management in a Presidential Campaign: A Case Study of the Obama Campaign's Management of Its Internet-Integrated Grassroots Effort. *Journal of Public Relations Research*, 313-335.

- Liben-Nowell, D., Novak, J., Kumar, R., Raghavan, P., & Tomkins, A. (2005). Geographic routing in social networks. *Proceedings of the National Academy of Sciences*, 11623-11628.
- Marx DSW, J. D. (2008). Online Fundraising in the Human Services. *Journal of Technology in Human Services*, 137-152.
- McSweeney, P. J. (2009). Gephi network statistics. Google Summer of Code, 1-8.
- Miller, B. (2009). Community fundraising 2.0—the future of fundraising in a networked society? *International Journal of Nonprofit and Voluntary Sector Marketing*, 365-370.
- Morgan, K. (2004). The exaggerated death of geography: learning, proximity and territorial innovation systems. *Journal of Economic Geography*, 3-21.
- Parlapiano, A., & Shorey, R. (2017, June 20). Who Financed the Georgia Sixth, the Most Expensive House Election Ever. Retrieved from The New York Times: <https://www.nytimes.com/interactive/2017/06/20/us/politics/georgia-6th-most-expensive-house-election.html>
- Perez-Truglia, R., & Cruces, G. (2017). Partisan Interactions: Evidence from a Field Experiment in the United States. *Journal of Political Economy*, 1208-1243.
- Prokop, A. (2016, February 11). Bernie Sander's remarkable small-dollar fundraising should scare Hillary Clinton. Retrieved from Vox: <https://www.vox.com/2016/2/11/10963902/bernie-sanders-fundraising>
- Quan, S., Zhang, T., Xu, C., & Hossain, M. (2015). Social Event Classification via Boosted Multimodal Supervised Latent Dirichlet Allocation. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 27.
- Raihani, N. J., & Smith, S. (2015). Competitive Helping in Online Giving. *Current Biology*, 1183-1186.
- Scellato, S., Mascolo, C., Musolesi, M., & Latora, V. (2010). Distance Matters: Geo-social Metrics for Online Social Networks. 3rd Conference on Online Social Networks (p. 8). USENIX Association. Retrieved from The Advanced Computing Systems Association: http://static.usenix.org/event/wosn/tech/full_papers/Scellato.pdf
- Sherfinski, D. (2017, February 21). Trump smashes Obama's small-dollar fundraising pace, bests Clinton, Sanders combined. Retrieved from The Washington Times: <https://www.washingtontimes.com/news/2017/feb/21/trump-smashes-obamas-small-donor-fundraising-pace/>
- Stein, S., & Cherkis, J. (2017, June 28). The Inside Story of How Bernie Sanders Became the Greatest Online Fundraiser in Political History. Retrieved from HuffPost: https://www.huffingtonpost.com/entry/bernie-sanders-fundraising_us_59527587e4b02734df2d92c1
- Valenzuela, S., Kim, Y., & Gil de Zúñiga, H. (2011). Social Networks that Matter: Exploring the Role of Political Discussion for Online Political Participation. *International Journal of Public Opinion Research*, 163-184.

- Vance, J. (2016). *Hillbilly elegy : a memoir of a family and culture in crisis*. New York: Harper Collins Publishers.
- Vesnic-Alujevic, L. (2012). Political participation and web 2.0 in Europe: A case study of Facebook. *Public Relations Review*, 466-470.
- Vogel, K. P. (2016, February 10). How Bernie bulk a fundraising juggernaut. Retrieved from Politico: <https://www.politico.com/story/2016/02/bernie-sanders-fundraising-219112>
- Wisconsin Ethics Commission. (n.d.). View Receipts. Retrieved from Wisconsin Campaign Finance Information System: <https://cfis.wi.gov/>
- Wright, G. C., Erikson, R. S., & McIver, J. P. (1985). Measuring State Partisanship and Ideology with Survey Data. *The Journal of Politics*, 469-489.
- Yano, T., & Smith, N. A. (2010). What's Worthy of Comment? Content and Comment Volume in Political Blogs. *ICWSM*.