University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Political Science Department -- Theses, Dissertations, and Student Scholarship

Political Science, Department of

Fall 11-1-2018

Preferential Politics

Cary Wolbers
University of Nebraska - Lincoln

Follow this and additional works at: http://digitalcommons.unl.edu/poliscitheses

Part of the Political Science Commons

Wolbers, Cary, "Preferential Politics" (2018). *Political Science Department -- Theses, Dissertations, and Student Scholarship*. 47.
http://digitalcommons.unl.edu/poliscitheses/47

This Article is brought to you for free and open access by the Political Science, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Political Science Department -- Theses, Dissertations, and Student Scholarship by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Preferential Politics

By: Cary Wolbers

A Dissertation

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements for the Degree of Doctor of Philosophy

Major: Political Science

Under the Supervision of Kevin Smith

Lincoln, Nebraska July, 2018

Table of Contents

Abstract 3

Chapter One 4

Chapter Two 17

Chapter Three 33

Chapter Four 49

Chapter Five 65

Chapter Six 79

Works Cited 88

Appendix A 94

Appendix B 101

Abstract:

Preferential voting is a unique system of voting that, while enjoying popularity abroad, has yet to make a significant impact on American political culture. However, within that past few years, preferential voting has been adopted by a number of cities across the country and the state of Maine. This dissertation examines the growing role of preferential voting in the United States, the impact of preferential voting on the electoral process, and the public's perception of preferential voting. This project uses survey data and data collected through Twitter to demonstrate that preferential voting is generally popular with the electorate and reduces campaign negativity, but it can confuse certain voters. Ultimately, this project demonstrates that preferential voting has the potential to address many of the complaints directed towards plurality voting.

Dissertation:

Chapter One: Introduction

Plurality voting is the primary institutional feature of elections at all levels of the American political system. Commonly referred to winner-take-all politics, in plurality systems voters can cast one vote for one candidate for each office on the ballot, and the candidate receiving the most votes wins. This system is such a ubiquitous institutional feature of American politics that it is rarely questioned, and alternative approaches to electing candidates to office are rarely considered. However, alternate approaches do exist, and they may offer notable advantages over plurality systems. Preferential voting is one such system. While preferential systems have many variants, their key contrast with plurality systems is that voters can vote for more than one candidate for any given office, thus allowing voters to cast a ballot that reflects a rank-ordered preference. This contrast between preferential systems and plurality systems is even more notable due to the increased number of candidates generally found in preferential elections (Bowler et al. 2003).

The primary purpose of this dissertation is to examine how preferential voting systems alter voter and candidate behavior, as well as voter attitudes towards electoral outcomes. This is a timely topic as a number of states and localities are considering adopting a preferential voting system, and such systems have already been implemented in a handful of places. According to the most recent data from Project FairVote, the cities of Oakland, San Francisco, San Leandro, and Berkeley (all in California) use some form of preferential voting to elect all of their city officials. Political parties in several states such as Utah and Virginia use preferential voting to nominate and select candidates

during the primary process. Arkansas, Louisiana, Mississippi, Alabama, and South Carolina allow preferential voting for military and overseas voters. Several other states and cities are considering preferential voting, considering ballot measures to implement preferential voting, or are awaiting the implementation of preferential voting. For example, Massachusetts and Maryland are currently considering bills that would implement preferential voting statewide. Maine recently became the first state to pass a citizen-initiated referendum that would adopt preferential voting in its statewide elections for governor and U.S. Senate. While Maine's governor has issued multiple legal challenges against the implementation of preferential voting, it appears that Maine's adoption of preferential voting is proceeding and has been cleared for use in the 2018 primaries. It should be noted that the Republican party in Maine has been opposed to the use of preferential voting (Mistler 2018). In June 2018, Maine voters became the first voters in the country to use preferential voting in a state primary. Governor LePage stated that he would "probably" not certify the results of the primary, although it is the secretary of state who certifies election results, not the governor (Nilson 2018).

Though small in number, these state and local government transitions from plurality to preferential voting systems represent a unique opportunity to answer the following questions: 1) Do voters prefer a preferential to a traditional plurality voting system? 2) What are the demographics of voters who are most likely to support a shift to a preferential system, and why do specific demographics prefer one voting type over another? 3) Can public support for preferential voting systems be reliably measured by using social media as a proxy for more traditional survey approaches to measuring public opinion? 4) Do preferential voting systems cause a shift in political campaign strategies?

Finding systematic answers to these questions is important for several reasons. First, the impact of preferential voting on electoral behavior and satisfaction with the outcome of those elections in the United States is virtually unknown. Given the historical rarity of preferential systems, there is unsurprisingly a paucity of existing research on this topic, though research conducted in other countries strongly implies that the impacts could be significant (see discussion next chapter). Second, because of the overwhelming prevalence of plurality voting in the United States, it has been extremely difficult to determine the extent to which the campaign process is influenced by the institutional structure of the voting process. This lack of variation has made it virtually impossible to empirically assess how the systemic context of elections shapes the behaviors of actors within those systems, even though there are good reasons to expect such behavioral effects exist. For example, there are good reasons to suspect that the implementation of preferential voting may encourage more candidates to run for office while simultaneously reducing the negativity of campaigns (see discussion below). Finally, pursuing these questions will help us understand how (or if) alternative forms of voting can impact overall voter satisfaction. If the analysis demonstrates that voters in preferential systems report consistently higher levels of political satisfaction, then it will provide evidence that such voting systems should be considered a viable, perhaps even preferable from a normative democratic perspective, alternative to plurality voting.

This project is not only valuable because it is addressed at filling a large gap in the existing research literature on the central mechanism of representative democracy in the United States. It will also make a unique and independent methodological contribution. This project includes the development and deployment of an original tool

for measuring public opinion consisting of a web scraper designed by Darren Wolbers (Master's degree student in computer science) and myself. The web scraper provides a way to collect and sort tweets as well as collect demographic and attitudinal information about the individuals who created the tweets. The web scraper provides a new social-media-based method by which to collect, measure and analyze public opinion on preferential voting (and potentially a large range of other issues). The use of Twitter data as a means of approximating public opinion is not a new concept; however, the web scraper designed for this project improves upon the process tremendously through the incorporation of more independent variables and a duplicate checking system. The sophisticated use of Twitter data presented in this analysis is what makes this methodological approach the first of its kind.

It is worth noting that Project FairVote has conducted several analyses using the same data as used in this dissertation. Socioeconomic and Demographic Perspectives on Ranked Choice Voting in the Bay Area (2015) was written by Caroline Tolbert and Sarah John was written for Project FairVote for the report was created for the purposes of providing broad analytic insights from the same survey used in this analysis. The report was not peer-reviewed. These analyses, while valuable, do not reflect the same depth of research and analysis presented in this research. For example, Project FairVote and I used the same 2014 survey data in analyses of campaign civility. While both projects share the basic similarities regarding the subject matter and the general conclusions (i.e. cities with preferential voting tend to have lower levels of perceived campaign negativity), the methods used to reach these conclusions are significantly different.

This project uses similar data and arrives at many of the same conclusions as the

2014 FairVote report. In fact, the FairVote report examines topics comparable to those which I address in Chapters Two, Four, and Five. However, the FairVote report analyses rely on descriptive statistics without the use of inferential analysis, and it also uses different variables. The fact that the FairVote report reaches many of the same conclusions as this dissertation while using such drastically different techniques is a testament to the robustness of the data provided in the 2014 FairVote Survey.

The dissertation project is thus centered on three major empirical contributions: first, it seeks to assess levels of public support for preferential voting and examine individual-level variation in that support. As will be discussed in-depth later, there are excellent theoretical and empirical reasons to hypothesize systematic differences in support for preferential voting among differing demographic groups. However, no extant scholarship has systematically identified the basic demographic variates of support for preferential voting in the United States. Some cross-national scholarship has attempted to tackle the question of support for preferential voting in a limited way (e.g. Poundstone 2008; Fazio and Gianluca 2014; Reilly 2001), but not to the extent detailed in this dissertation. Second, this analysis represents the first comprehensive empirical examination of systematic differences in campaign satisfaction, candidate satisfaction, and campaign methods between plurality and preferential voting systems in the United States. Since preferential voting is quite new in this country, no previous research has drawn definitive conclusions about the comparative impact of the two voting systems in United States on variables such as public opinion, voter behavior, or candidate behavior. Finally, this analysis makes a significant methodological contribution through the use of the proprietary web scraper program designed to collect and analyze tweets. This

program provides a new tool to approximate public opinion and gathering data on groups that may not be represented through traditional surveys. While previous scholarship has only used tweets as a proximal measure of public opinion, the program created for this analysis also collects key demographic characteristics such as age, gender, and ethnicity. Thus, this tool analyzes not only what people are tweeting but provides demographic data on who is tweeting it.

This project is organized around three core research questions that encapsulate the intended contributions as described. Specifically, these research questions and a basic synopsis of how this dissertation project will seek to answer them are as follows:

Research Question One: *"What demographic characteristics predict support for preferential voting and why?"*

To address this first research question, the dissertation will test the hypotheses that young respondents, low-income respondents, female respondents, and minority respondents are more likely to support preferential systems. These hypotheses will be empirically tested in Chapter Two using survey data and in Chapter Three using data collected from Twitter. It should be noted that income will not be tested in Chapter Three, as there was no accurate way to collect this information from Twitter users. These traits were selected for hypothesis testing because they are well known to correlate with lower levels of political participation and/or historical political disenfranchisement and this may at least be partially due to the institutional arrangements of plurality voting systems. Plurality voting forces these groups to choose between a more limited field of candidates, and as a result, these groups are less likely to engage in the political process. Younger voters, for example, are not as entrenched in the system of plurality voting as

older voters, making them more receptive to the idea of a different system. A voter who is sixty or seventy years typically has decades of experience with plurality voting. That older voter has a greater familiarity with the system and how it works than a twenty-year-old with much more limited experience of going to the polls. As a result, the older voter may have a greater stake in keeping the status quo. Conversely, the younger voter has much less invested in the existing system and may be more willing to try something new.

Similarly, voters who may have legitimate frustrations with the existing system may be more open to preferential voting. Supporters of third parties and traditionally marginalized voters such as racial and ethnic minorities and low-income voters may be highly familiar with the existing plurality system, but due to their level of dissatisfaction with their elected officials, they have reason to want to try a preferential system. The preferential voting process will give them more options to express their voices and potentially make the political system more responsive to their concerns. Marginalized voting groups, especially racial groups, are also more likely to support preferential systems because they largely eliminate, or at least greatly reduce, the practice of political outbidding, which can occur in plurality voting systems. Political outbidding is the practice of promising policies that disproportionately benefit a particular group to curry electoral favor with that group (see next chapter for a full discussion).

Finally, individuals who advocate for preferential voting systems are predicted to be better educated. This is predicated on the assumption that those who are in favor of preferential processes must not only have some understanding of the potential downsides of plurality systems, but also some understanding or at least knowledge of alternative voting systems. In short, those who support preferential systems are hypothesized to be

higher information voters, at least within this domain, and that connotes higher levels of education.

As a means of addressing common criticisms levied towards preferential voting, this dissertation will also examine whether minority voters, low-income voters, female voters, and younger voters have more difficulty understanding either the concept of preferential voting or the instructions associated with engaging in the preferential voting process. While preferential voting is more complicated than plurality voting, this analysis will demonstrate that the difference in the level of complexity between plurality voting and preferential voting is not sufficient to cause a statistically measurable effect, a conclusion that is supported by the widespread adoption of preferential voting worldwide.

Research Question Two: *Can public attitudes toward preferential voting be accurately measured using data gathered from social media posts?*

Social media is a prominent contemporary means by which individuals freely express their support or opposition to policies, politicians, or political groups. As a result, social media provides a potentially rich source of data regarding voter attitudes and behavior. This was crucial to my project since there has been very little data collected on opinions and behavior relevant to preferential voting systems in the United States. Maine is the only state that has subjected a preferential system to a statewide ballot initiative and, as is typical in the contemporary political environment, this was an issue discussed extensively on social media platforms. Thus, social media activity linked to the Maine initiative offered an excellent opportunity source of data for political attitudes on preferential voting. Effectively this offers a way to potentially measure and analyze

public attitudes on preferential voting even though traditional survey data on this issue is largely non-existent.

The huge challenge in mining this potential data source was developing a software program capable of collecting and analyzing social media posts in a way that produces data reasonably comparable to that collected by traditional survey methods. A central contribution of this dissertation thus describes our success in developing a tool that overcame this challenge, and this project demonstrates its capabilities through a comprehensive analysis of attitudes through tweets made during the Maine referendum. The web scraper provides a means of collecting data that can be used to test hypotheses about preferential voting in a manner comparable to using survey data. It assesses the demographic characteristics of Twitter users, attaches that information with the users' corresponding tweets, and rates the emotional polarity (positive or negative) of their tweets. Finally, to provide additional insight into the demographics of people who are likely to support a shift from the status quo, the web scraper determines the reading level of each tweet.

If successful, these data collected through the webscraper should support an analysis that replicates the findings of Research Question One. In doing so, this project will demonstrate the usefulness of the analytical tool as an alternative means of obtaining data on public opinion.

Research Question Three: *Do preferential voting systems cause a shift in political campaign strategies?*

The type of electoral system affects political candidates' incentive to engage in negative campaigning as a means of attracting voters. Under plurality voting, a voter is

given a single choice regarding each candidate. Candidates thus compete in a zero-sum context where any vote cast for one candidate is a loss for the opposition. Thus, candidates in plurality voting systems have strong incentives to criticize and attack their opponents as a means of setting themselves apart from their competition and demoralizing their opponent's supporters.

Preferential voting systems alter the nature of electoral campaigns by fundamentally changing the incentives that candidates have for attacking their rivals (Donovan et al. 2016). Voters can cast their votes for multiple candidates, so candidates do not need to engage in a winner-take-all campaign strategy. Preferential systems allow candidates to fight for a voter's second or third preference, and those second or third preferences can have an important impact on the electoral income. Trying to win a voters' second or third preference by trying to tear down the candidate who is their first preference is, at best, a risky strategy. A candidate has clear incentives to position themselves as a viable option even for voters who do not rank them first.

Preferential voting systems may also apply downward pressure on campaign negativity by widening the field of candidates, thus reducing the possibilities of a binary, us-against-them choice in a general election. More candidates may dilute the impact and visibility of negative campaigning simply because there may be more candidates who would need to be attacked. Instead, preferential voting systems passively encourage candidates to focus on voter mobilization instead of attacks against his/her political opponents. As a result, cities with preferential voting report reduced perceptions of campaign negativity (Donovan et al. 2016).

Research Question Four: *Are voters capable of understanding preferential voting?*

The adoption of preferential voting is not without controversy. While some politicians may be concerned that preferential voting will to cause large shifts in electoral power, many journalists, pundits, and voters have expressed concerns regarding the electorate's ability to understand and engage with preferential voting because it represents a significant departure from the plurality system with which most voters are familiar.

There are significant concerns that the general public may be unable to understand the complexities of preferential voting (Weil 2016). These concerns largely stem from the increased complexity of preferential voting ballots requiring that voters learn about a greater field of candidates (Neely et al. 2005; Cook & Latterman 2011; Arrow & Raynaud 1986). Additionally, some voters may not be able to understand the instructions they are required to read prior to voting (Weil 2016) These concerns are not unwarranted and demand further examination. Chapter five will examine whether there are specific socio-demographic groups who experience increased difficulty understanding preferential voting or the instructions they are required to read prior to engaging in the preferential voting process.

Conclusion

The discussion above introduces the main research questions to be addressed in this dissertation, provides an argument for why answers to those research questions are worth pursuing and delivers a brief summary of the analytic approach that will be employed to answer those questions. The rest of the dissertation will proceed as follows:

Chapter Outline:

Chapter 2: Research Question One

Using traditional city-level survey data, this section will primarily focus on voters' support and satisfaction with preferential voting and how demographic characteristics predict support for preferential voting. Specifically, this chapter will focus on whether voters who are frequently subjected to political outbidding tend to support preferential voting more than other demographics.

Chapter 3: Research Question Two

This chapter will conduct a comparable investigation to that reported in Chapter Two, i.e. it will investigate support for preferential voting and test theoretical expectations about variation in that support by demographic groups. This analysis, however, will use data on a state (Maine) referendum initiative gathered from social media. In doing so, the discussion in Research Question Two will introduce a new web scraper tool for collecting and analyzing public opinion data.

Chapter 4: Question Three

This chapter will analyze the differences in campaign strategies encouraged by the two voting systems by examining differences in campaign tactics and activities between cities with preferential voting and cities with plurality voting. This comparison will be used to draw conclusions regarding the potential shifts in campaign strategies that accompany the transition from plurality voting to preferential voting.

Chapter 5: Understanding Preferential Voting

One of the common criticisms of preferential voting is that it is more complicated than plurality voting. Concerns that voters cannot understand how a preferential system works raise questions about whether this system should be more broadly adopted,

regardless of whether voters support such a move. Accordingly, this chapter focuses on examining voters' understanding of the instructions they are required to read prior to voting in a preferential election. This will provide evidence either supporting or disproving the legitimacy of many of the criticisms levied at preferential voting, specifically that the voters do not understand how to participate in the voting process.

The second goal of this chapter is to examine whether respondents have a difficult time understanding preferential voting systems – or do they have difficulty understanding voting systems generally.

Chapter 6: Conclusion

Chapter 6 summarizes the findings and inferences that can be taken from the analyses presented in the dissertation and discusses their implications. It also discusses the pros and cons of preferential systems and ponders the future of preferential voting in the United States and whether broader adoption of such systems is justified.

Chapter Two

The primary objective of this chapter is to isolate and identify the socio-demographic factors which predict support for preferential voting, i.e. it is focused on addressing the first research question listed in the previous chapter. This will be accomplished using one of the few datasets that include extensive information about attitudes on plurality systems in the United States. This data was collected from subjects in various cities in California, some which use a preferential voting system and some which use a plurality system. For theoretical reasons described below, I hypothesize that voters' age, income, ethnicity, and gender will be significant predictors of support for preferential voting in the United States.

Literature and Background

There are good reasons why voters, especially certain groups of voters, would be motivated to support shifting from a plurality voting system to a preferential voting system. To understand why it is important clearly define the two general systems under discussion and clarify the key arguments made in favor of shifting to a preferential system. Plurality voting is generally defined as a competition between candidates in which each member of the electorate casts a single vote for a single candidate (Poundstone 2008). In plurality voting, only one candidate can win, which is why this system is often referred to as a winner-take-all or a first-past-the-post (Poundstone 2008) contest. Within the context of the United States, voters are typically given the ability to vote for one of two candidates who represent the political platforms of the Democratic or Republican Party. There are mechanisms in all states for third-party and independent

candidates to gain access to the ballot—and plurality contests do sometimes include non-major party candidates—but ballot access for such candidates is far from guaranteed, and even individuals who negotiate the challenge of actually getting on the ballot rarely have the mainstream appeal to gain elected office. In the typical election in the United States, each major party puts forth a candidate, and voters must decide which of these candidates deserves their vote. Even in putatively non-partisan elections at the local level, candidates often openly align with (and are backed by) one of the major political parties. Regardless of partisan considerations, the winning candidate in any plurality election is simply the one who receives a majority of the votes.

In contrast, preferential voting allows members of the electorate to vote for more than one candidate, ranking their selected candidates based on preference (Obata & Ishii 2003). In a preferential voting election, if a majority of voters (i.e. > 50 percent) select a particular candidate as their first preference, the election is over. However, if a candidate fails to garner enough first preference votes to achieve a majority of first preference votes an elimination process begins. Candidates are eliminated if they do not have a sufficient number of voters' first choice votes, and they are eliminated based on their respective share of the overall vote until a single candidate remains (Obata & Ishii 2003). For example, consider an election where there are five candidates. When voters cast ballots, they rank the top three candidates as 1, 2, and 3 in order of their preference. When the votes are tallied, if one candidate receives an outright majority of 1s, in other words, they are the first preference of at least half the voters casting ballots, that candidate wins. If no candidate meets that threshold, the candidate with the lowest number of 1s, i.e. the candidate with the lowest number of first preference votes, is eliminated. The second

preference votes of those who ranked the eliminated candidate first are now counted for the four remaining candidates. This process iterates until one candidate has an absolute majority. Thus, the key difference between preferential versus plurality voting is that preferential voting does not instantly negate a voter's satisfaction if his or her first preference is unsuccessful. Voters may see their second or third choice candidates achieve electoral success, affording them at least some measure of satisfaction with the election results, even if their most preferred candidate is not elected. The discussion below details the advantages of preferential voting which appeal to the general electorate.

Preferential voting minimizes the electorate's need to engage in strategic voting (Bartholdi & Orlin 1991). Strategic voting occurs when a voter supports a candidate, who they may not sincerely support, in order to prevent an undesirable electoral outcome (Farquharson 1969). For example, during the 2016 election voters who supported the Green Party's candidate, Jill Stein might instead have chosen to cast a ballot for Democrat Hillary Clinton, a choice incentivized by the low probability of third-party success in a winner-take-all system. These voters, in short, are effectively being forced to choose between voting for a Republican or Democrat who may not fully represent their political preferences and voting for a third-party candidate (such as Ms. Stein) who may represent their true political preferences but doesn't have a realistic chance of winning. Preferential voting minimizes the need to engage in the strategic voting incentivized in plurality voting by giving voters the ability to vote for both: the candidate who represents their political preferences and the candidate who is most likely to win. For example, the voter who supported Jill Stein could have selected Stein as her first choice, kept Hillary Clinton as her second choice, and selected Jeb Bush in as her third choice.

Research has generally found that, compared to plurality systems, preferential voting tends to have a positive impact on voter engagement in political campaigns, promoting higher levels of voter mobilization, and thus producing higher levels of turnout (Bowler et al. 2003). These findings have important implications for well-known and repeatedly demonstrated concerns about the low voter turnout common with the plurality system in the United States. It is widely accepted that reducing political engagement threatens democracy's ability to function properly (Polsby 1963), and an extensive literature in political science suggests that any blame for a failure to broadly engage in the political process lies squarely with the electorate, who lack the interest, engagement, or knowledge to effectively engage in the political process (e.g. Wolfinger & Rosenstone 1980).

There are strong reasons to suspect that a lack of political participation cannot simply be assigned to civic-shirking by voters. There are also systemic reasons for downward pressure on turnout. These include the overall unpopularity of and declining trust in American political institutions (Hibbing & Theiss-Morse 2002). Voters who do not trust a government and do not believe it can functionally represent them are unlikely to participate (Hetherington 2005; Hetherington & Rudolph 2015). A preferential system may, at least partially, be able to reverse or at least lesson such institutional causes of distrust and lack of participation. Dissatisfaction with electoral system's status quo and the noted deterrents to political engagement cannot, of course, be rectified by simply adopting a preferential voting system. There are good reasons, however, to believe such a shift could significantly ameliorate them. Existing research already indicates that not only do voters tend to readily grasp the potential advantages of a preferential system, but they

will also support such a system if the choice is offered, and once in place, such systems will actually produce some of the benefits promised. Bowler, Brockington, and Donovan's study on alternative voting styles (2003), for example, found that cumulative voting, which is a type of preferential voting, has the potential to counteract the behavioral and institutional barriers to voting and typically results in a small increase in voter turnout.

The existing literature makes several arguments about why preferential systems not only appeal to the electorate at large but also appeal especially to certain groups of voters that historically have lower participation rates. First, existing scholarship suggests that such systems offer obvious benefits to traditionally disenfranchised groups, such as racial and ethnic minority voters, female voters, and young voters. An important reason for this is because preferential voting discourages political outbidding, which is the practice of promising policies that disproportionately benefit a particular group to curry favor with that group. This practice may target minority groups, ideological groups, religious groups, or ethnic groups (Fazio and Gianluca 2014). For example, the Republican Party has recently engaged in a prominent example of political outbidding as it proposed policies that appeal to evangelical Christians and predominantly white, anti-immigrant voters. Such outbidding is not limited to one partisan or ideological group.

The Democratic Party has traditionally supported policies such as affirmative action that promise to directly benefit racial and ethnic minorities, and these policies are at least perceived by some as being at the expense of whites. The bottom line is that it is politically beneficial to propose policies that engender support from such an influential voting demographic even if doing so creates division and conflicts with other social

groups and can exacerbate existing political and ethnic rifts (Kanchan 2005).

While candidates and political parties can and do engage in outbidding in ways that seek to benefit traditionally disenfranchised groups, outbidding tends to disproportionately favor majority groups for the simple reason that those groups constitute the biggest blocs of voters. Outbidding to such groups simply provides the candidate with a higher chance of reaping an electoral return. Accordingly, it is not surprising that traditionally disenfranchised groups, especially racial minorities, tend to favor preferential voting (Fraenkel & Grofman 2006). Such support is justified by empirical research that suggests not only do preferential systems reduce the incentives to engage in classic outbidding, they also explicitly encourage a more cross-ethnic approach to campaigning (Reilly 2001; Neely et al. 2005). Preferential systems reduce the need to engage in political outbidding because of the increased political choice preferential voting affords the electorate. Plurality voting forces voters to choose a single candidate, but preferential voting allows voters to select multiple candidates. Candidates under a system of plurality voting often neglect to reach out to areas or groups who they perceive as definitively going to their opposition; however, under preferential voting these candidates still engage with these groups as being a second or third choice vote can still lead a candidate to victory (Reilly 2001; Neely et al. 2005).

For similar reasons, preferential voting may also appeal to female voters. Research suggests that preferential voting encourages political moderation from candidates and greater gender diversity in government. Under a system of preferential voting, women have better chances of attaining elected office because voters are not forced to decide between voting for a woman and voting with their ideological leanings

(McElroy & Marsh 2010; Cook & Letterman 2011; Neely & Bash 2005). Preferential voting may also appeal to female voters in that the transition from plurality voting to preferential voting is often accompanied by a reduction in campaign negativity, which will be empirically demonstrated in Chapter Four. Female voters generally dislike negative campaigning more than their male counterparts and rely on negative campaigning less when running for elected office (Herrnson & Lucas 2006; Herrson et al. 2003). A system which reduces the prevalence and necessity of negative campaigning thus should appeal to female voters and potential female candidates. Female voters, just like other groups who have historically been politically disenfranchised, are likely to benefit from preferential voting because it would reduce the political outbidding that is often directed towards them (Reilly 2001).

A plurality system may also help drive up participation among younger voters, a notoriously low-turnout group in the United States. Partially this is because younger voters tend to be more accepting of political change than their older counterparts (Maccoby 1954; Earl & Kimport 2011). Younger voters are more likely to support alternative forms of voting, specifically preferential voting because they are less entrenched in plurality voting as the accepted political norm (Haan et al. 1968). Younger voters are more likely to support an alternative voting system which promises improvements to political efficiency and efficacy while older voters are more likely to be content with a system to which they are acquainted.

As with younger voters and female voters, preferential voting may appeal to low-income voters (Rabushka & Shepsle 1972; Bartels 2008). Low-income voters, like minority voters, are frequently subjected to political outbidding, wherein politicians seek

to curry favor among members of a group by proposing policies which will help that group more than the policies proposed by the politician's opponents. Candidates are frequently vying for the support of the economically disadvantages by promising various policies that will help them improve their financial situations. Ultimately, these voters may be placed in a position where they are forced to choose between a candidate who represents their political beliefs and ideologies and another who may offer beneficial economic policies. Preferential voting allows low-income voters to avoid the trap of political outbidding by granting voters greater political options at the ballot box.

While there is little empirical data available, it also possible that there will be partisan differences in support for plurality systems. The socio-demographic groups just described tend to be core constituencies of the Democratic Party, so it makes sense that this would translate into greater levels of support for preferential systems among Democrats. Also supporting this hypothesis, however, are long-standing ideological differences between the two major political parties in the United States. The Republican Party is more associated with conservatism, i.e. a reluctance to change the traditional status quo, and on top of that, there may be rational reasons for GOP voters to support a shift to a preferential system that may boost the electoral opportunities of more Democratic-leaning candidates drawn from minority, low SES and female candidates. Democratic partisans, on the other hand, have rational reasons to support a shift to a plurality system for exactly those reasons.

The discussion above suggests that several aspects of plurality voting will make it attractive to voters generally, and more specifically to particular demographic groups that have arguably been disadvantaged by the traditional plurality system, and have much to

gain by a shift from that status quo. The remainder of this chapter will test a series of hypotheses drawn from these theoretical expectations, analyzing how demographic characteristics map onto support for preferential voting. The specific hypotheses to be tested are as follows:

Hypothesis One: Support for preferential voting will decline with age.

Hypothesis Two: Low-income voters will be more likely to support preferential voting than their middle-income or high-income counterparts.

Hypothesis Three: Minority voters will be more likely to support preferential voting than their white counterparts.

Hypothesis Four: Female voters will be more likely to support preferential voting than their male counterparts.

Methods and Data:

To test these hypotheses, I use data collected through the Rutgers-Eagleton Poll, which is carried out through Rutgers University's Center for Public Interest Polling. Established in 1971, the Eagleton Center for Public Interest Polling is one of the country's most respected university-based research centers. The goal of the surveys was to understand the impact of preferential voting, as well as those who support it. Conducted in 2014, this survey was carried out in eleven cities in California and included 2,400 participants. The analyses contained in Chapter Two will only be utilizing the data from this survey.

The 2014 survey was conducted in both cities with preferential voting and those without. The samples from each city are proportional to their population with the smallest samples of 100 respondents coming from Anaheim and Santa Ana and the largest sample

of 685 coming from Oakland. These cities are a combination of cities which either have ratified measures to allow preferential voting at the local level and cities who had, at the time of collection, retained plurality voting for local elections. Cities without preferential voting were chosen based on population, socioeconomic factors, and ideological composition to provide a viable counterpart to the cities with preferential voting. To ensure that no cities were disproportionately represented, the survey implemented strict respondent quotas based on each city's population. Respondents in both types of cities were asked to answer the following question:

"*Do you think ranked-choice voting, where voters can rank candidates in order of preference with their first choice counting most, should be used in local elections in your city*?

This question was used to create the binary dependent variable for the logistic regressions that will be used to test the above hypothesis. Respondents who responded with "yes" were coded as *one*, while all other respondents were coded as a *zero*. The use of this question as the primary dependent variable allows for a valuable proximal measure of respondent's perceptions of preferential voting.

The key independent variables in this analysis are a series of dummy variables for race, gender, and income. Age was measured with an ordinal variable. In this analysis, we were testing the support for preferential voting within California's white and nonwhite populations; therefore, this analysis coded race in terms of white respondents (coded 1) and nonwhite respondents (coded 0). Income was tested using two different approaches. First, income was measured through a series of dummy variables representing high-income respondents, low-income respondents, and a dummy variable

for respondents who refused to specify their income. Respondents who chose not to specify their income were dummy coded as a means of incorporating them into the analysis without imputing their income. Excluding the respondents from the analysis would damage the external validity of the analysis and imputing income for these respondents could damage the accuracy of the analysis. This analysis tests both this coding strategy and imputation to test income in multiple ways. This strategy has been utilized by multiple scholars and is an accepted means of dealing with missing data (Bhat 1994). Respondents who specified making under $50,000 annually were placed in the low- income category, while respondents who indicated earning over $75,000 annually were placed in the high-income category. Respondents who earned between $50,000 and $75,000 annually were chosen as a reference category and were excluded from the analysis. The second measure of income was created by imputing missing income variables through Tobit regression. The Tobit imputation model included age, gender, ethnicity, and education as explanatory variables. Generally speaking, these Age, was coded as an ordinal scale based primarily on decadal increments from 18 to 99 years old.

This analysis used two separate measures of partisanship. First, the analysis used an ordinal measure of partisanship. Partisanship was controlled for to ensure that the analysis accounted for hostility towards preferential voting that may be the result of partisan leaning. As demonstrated by Maine's governor Paul LePage, there can be serious partisan-based hostility towards preferential voting; therefore, controlling for partisanship ensures that any of those hostilities are accounted for. This measure of party identification was organized along a three-point ordinal scale ranging from respondents who identified as "Democrats" to respondents who self-identified as "Republicans."

Second, a series of dummy variables were created to represent respondents who specify a particular partisan affiliation. Respondents who identified a strong partisan affiliation were categorized as a partisan. Conversely, respondents who identified as independent, or an independent with a particular partisan leaning, were coded as independents.

The analysis also controlled for education and employment status. Education was coded on an ordinal scale ranging from respondents who had not completed high school to those who had completed college or graduate school. Employment, like income, was included to control for various economic factors that might influence a respondent's ability to engage with his/her respective voting system. For example, employed respondents may not appreciate the increased time they are required to spend learning about multiple candidates in a preferential election. Finally, I included a control variable for the type of city a respondent was from. This variable was dummy coded with respondents from cities with preferential voting being coded as one and respondents from cities with plurality voting being coded as zero.

This analysis will test the key hypotheses utilize several statistical models. Four separate logistic regression models will be used to test how specific demographic characteristics influence the respondent's attitudes towards preferential voting. The conclusions reached in this chapter will be used and built upon throughout the course of this dissertation. Therefore, it is important to ensure that the conclusions reached in this chapter hold up across different models using different coding strategies. The four models used in this analysis vary in their coding of partisanship and income. Income is one of this chapter's key independent variables, so testing different coding strategies is important to ensure the robustness of the result. Additionally, as previously explained,

preferential voting undermines the strength of political parties through increased political options. Testing the impact of partisanship through binary and ordinal categories allows me to test if there is a significant difference between Republicans and Democrats and partisans and independents.

Findings:

Table 2.1 (Appendix B) reports some basic descriptive statistics. While these tables do not represent the predictive findings, they do help establish a basis of comparison between cities with preferential voting and cities without preferential voting. This table shows a combination of theoretically relevant statistics and statistics which are used as control variables. These control variables were included in Table 2.1 to show that the sample is fairly representative of the general population. Additionally, this table shows basic descriptive statistics pertaining to the respondent's perceptions of various aspects of preferential voting. This table demonstrates a few key findings. For example, most respondents who live in cities with preferential voting believe that preferential voting should be used in local elections. While voting statistics tell us that this number likely does not reflect the true proportion of voters who actually participated in the electoral process, it is interesting to note that this number is higher than most exaggerated claims of voter participation; therefore, it is reasonable to assume that this preferential voting population was disproportionately politically engaged.

Table 2.2 (Appendix B) presents the findings of the statistical models. This model tests all four of the primary hypotheses. These findings support the hypotheses pertaining to inclusion, that historically disenfranchised groups were more likely to support preferential voting. Respondents who belong to groups commonly subjected to political

outbidding were more likely to support the continued implementation of preferential voting. Specifically, younger voters, minority voters, and lower-income voters were all more likely to support the continued implementation of preferential voting. Age was a particularly significant predictor. The results presented in Table 2.2 offer strong evidence for Hypothesis One. In all models, age is consistently, negatively related to support for preferential voting. A younger respondent is significantly more likely to support preferential voting than an older respondent. Across all models, female respondents were less likely to support preferential voting than their male counterparts. This finding was unexpected and is in the opposite direction to that hypothesized.

Given that there is no clear theoretical reason for why females, compared to males, would oppose preferential voting, the obvious explanation for this finding is that the model excludes some important gender-based interaction effect (e.g. a marginal impact for being white and female or highly educated and female that once accounted for would change the sign of the independent variable for gender). I attempted to account for such a relationship a number of ways; however, there were no significant differences between male and female respondents with regard to education and ethnicity. Neither did the interaction effects for party affiliation and income yield significant results despite there being noteworthy differences between male and female respondents regarding their partisan identities and their incomes. Regardless of measurement and model specification, the results consistently indicated that females are less likely to support preferential voting systems, though these failed to provide any indication of what that relationship is negative. These findings demonstrate that there may be detriments for women in cities with preferential voting not captured by this particular survey.

Ultimately, the findings presented in Table 2.2 lend strong support for Hypothesis One, Hypothesis Two, and Hypothesis Three.

Voters who are younger, low income and non-white are consistently found to be more likely to support preferential voting. These results do not lend support to Hypothesis Four. While I suspected that female respondents would prefer preferential voting, female respondents were statistically less likely to support preferential voting than male respondents.

Regarding partisanship, results indicate Republicans dislike preferential voting compared to nonpartisans. Conversely, Democrats showed a positive relationship with preferential voting compared to their nonpartisan counterparts, but these results were not significant. The significance of these findings only extended to the binary variables measuring party identification. Results for the ordinal measure of party identity were not significant. As previously stated, preferential voting often allows voters to vote for multiple candidates from across the political spectrum; therefore, preferential voting is more beneficial to independents than to partisans, which is reflected in these results.

Conclusion:

The primary goal of this chapter was to establish a basic set of theoretical expectations which might explain differences in support for preferential voting systems. This led to a series of hypotheses on the likelihood of support for preferential voting among specific demographic groups. At least in the California cities included in the survey data, the willingness of Americans to support the implementation of preferential voting was clearly dependent on several key factors that align with my hypotheses. First, the demographic characteristics of a respondent were highly predictive in determining

that individual's willingness to embrace preferential voting. Voters who typically are subjected to political outbidding – i.e. minorities, low income – were more likely to support preferential voting. This finding indicates that preferential voting may be the key to greater political involvement by minority and low-income voters. Age was also a significant factor in determining a favorability towards preferential voting; however, this could be understood through the inclusion hypothesis; younger voters and female voters are still frequently subjected to political outbidding.

In summary, the key takeaway from this analysis is that minority voters, low-income voters, and younger voters are all more likely to support the implementation, or continued implementation, of preferential voting. Female voters were less likely to support preferential voting. These groups typically have the most to gain from the ability to vote for multiple candidates. In the upcoming chapter, I will address the limitations of the survey data by utilizing an innovative methodological solution which serves as a source of new information and a robustness check on the conclusions found in this chapter.

Chapter Three:

The findings in the previous chapter offer mixed support key theoretical expectations, but they rest on a single survey done in a single state, a limitation that raises legitimate questions about the generalizability of the inferences drawn from the analysis. It would be optimal to assess the robustness of these findings by repeating similar tests using data drawn from multiple samples, ideally at the state level, because it is there that the constitutional authority to alter, change, and adopt new voting systems resides. Unfortunately, such data sources simply do not exist due to the scarcity of preferential systems in the United States. This may change in the near future as more jurisdictions begin seriously considering or actually adopting plurality systems, but currently, there simply is a general lack of data on attitudes on this issue.

To address this lack of data, this chapter utilizes a new method for data collection and employs it to analyze public attitudes surrounding Maine's Question Five, that state's 2016 ballot initiative to establish a preferential voting system. In doing so, this chapter addresses two primary objectives. First, it tests a subset of the same basic hypotheses tested in the previous chapter using a new source of data collected in a different state. This constitutes an attempt to assess the generalizability of the findings just reported. Second, and perhaps even more importantly, in the long run, this chapter introduces an entirely new tool with which to collect and analyze public opinion data.

Rather than traditional survey data, the method utilized in this chapter seeks to systematically analyze public opinion using Twitter as a data source. There are several reasons for taking this approach. First and most obvious is the lack of existing traditional survey data. As previously mentioned, preferential voting remains a relatively small

political phenomenon within the United States; consequently, major polling agencies have not dedicated significant resources towards studying it. Furthermore, given the lack of attention to preferential voting by American politics scholars, few researchers have dedicated their own money towards fielding surveys to promote a greater understanding of preferential voting in the United States (as far as I am aware, the survey used in the previous chapter is the only data source capable of supporting empirical analyses on preferential voting that are available to me). This means assessing the validity or generalizability of the findings presented in the previous chapter leaves me with little choice but to seek alternate data sources.

This effort is also motivated by a desire to assess Twitter data as a practical alternative to, or at least a valuable supplement to, traditional survey methods. Finding an efficient way to collect attitudinal data is one of the most consistent and perplexing challenges facing social scientists, Social media platforms such as Twitter—where people provide voluntary, unfiltered views—clearly has potential to address this challenge. I sought to construct exactly such an analytical tool premised on marrying the fields of computer science and public opinion research. In doing so, I built on previous scholarship and sought to extend existing efforts along this line by improving by collection and sophistication of the type of attitudinal data collected.

Specifically, this tool is a computer program that collects tweets, performs a sentiment analysis on those collected tweets, and extracts socio-demographic information pertaining to the users who created the tweets. A sentiment analysis is a process by which the emotional polarity of an opinion is identified based on the words present in that opinion. In this context, sentiment analysis essentially boils down to a set of algorithms

designed to identify and quantify the affective valence of a written statement. Sentiment analysis is typically used to better understand the writer's opinion towards the item, person, or place he or she is writing about. A sentiment can be positive, negative, or neutral (Pang & Lee 2008). By using a sentiment analysis on tweets, the program can rapidly evaluate a large number of people's opinions towards preferential voting, i.e. whether they are positive or negative. Thus, the program can provide a comparable dependent variable to those used in the analyses in the previous chapter.

The program is also capable of generating socio-demographic independent variables similar to those used in the previous chapter to test key hypotheses. It does this by using machine learning and computer vision to identify faces of the Twitter users and the demographic characteristics associated with those faces. A convenient feature of this program allows data to be exported directly into a CSV file, which ensures that the data is compatible with nearly every statistical package. The exported file looks very similar to a spreadsheet that might accompany a traditional survey, with each row representing a different user and each column representing a different variable. In short, the program produces a dataset stacked very similarly to a traditional survey, with rows representing subjects and columns variables (for the analysis conducted in this chapter the columns consist of positive/negative affect toward plurality voting, and age, gender, and race). The program is also designed to filter out individual opinions from excerpts tweeted from news stories or media reports. It does this by creating a measure of the percent of the text in the tweet that is found within closed quotation marks. This can be used in the data cleaning process to eliminate tweets which may not reflect a user's views on a particular topic but may instead reflect another user's views or an outlet's views on a particular

topic. For example, a tweet containing a quote by Maine's Governor Paul LePage about preferential voting is not useful for this analysis, as it does not reflect the true opinion of the person posting it. Any tweet which had more than 15% of the tweet in closed quotation marks was excluded from the analysis.

Utilizing Twitter as an Alternative to Survey Research:

Before employing this new tool to gather and analyze attitudes, however, is it is important to examine social media posts to assess whether they really represent a viable alternative to data collected from traditional survey methods. Survey research has largely dominated the study of political behavior and public opinion for good reasons. Scientific surveys have been the gold standard for measuring public opinion for decades. Besides the lack of data challenge faced in this dissertation project, what are the more general arguments for trying to extract public opinion data from social media? One such argument centers on the limitations of traditional survey methods.

Certainly, the use of scientific surveys has improved the discipline's understanding of how individuals shape, form, and change their opinions. However, survey research is not without its drawbacks. Traditional surveys are incredibly expensive. A representative survey may cost thousands of dollars, or even tens of thousands of dollars, depending on the target population and the length of the survey.

Second, a truly representative sample is becoming increasingly difficult to acquire. To conduct a representative phone survey requires a sample of phone numbers. While landline numbers are the easiest and cheapest phone numbers to acquire, landlines are becoming obsolete, particularly among younger Americans. Cell phones numbers provide access to a greater swath of the population (Link et al. 2007). But these phone

numbers can be extremely expensive to obtain. These issues are further compounded by the common issue of refusals or nonresponsive participants (Link et al. 2007; Dillman et al. 2014; Fowler 2013). Telephone surveys only have a response rate of around 18% (Kaplowitz et al. 2004). It can be extremely difficult to properly field a representative phone survey in due to the costs associated with acquiring a representative collection of phone numbers and the difficulty in finding respondents who are willing to take the survey. Additionally, survey research can be slow. It may take weeks or even months to properly field a survey, and this undermines the survey's ability to capture respondents' opinions towards timely political issues. The program used in this analysis has the ability to return a tremendous amount of data quickly and inexpensively. Therefore, it is valuable to evaluate this program's utility to supplement traditional survey research.

Online surveys have helped to address the issue of slow turnaround associated with traditional survey methods; larger online survey companies like YouGov and Qualtrics rely on representative panels giving them available and easily accessible audiences for the surveys. While leveraging the resources of these online surveys is cheaper than traditional phone surveys, it still cost thousands of dollars. There are free and inexpensive online survey options available, such as SurveyMonkey, Mturk, and Microworkers. While these options address the question of cost like traditional survey methods, they are hampered by low response rates. Ultimately, while online surveys provide an alternative to traditional survey methods, they still must contend with many of the same obstacles.

There are advantages and disadvantages to trying to address some of the challenges in traditional survey approaches by mining social media like Twitter. Perhaps

the most notable and obvious is that Twitter data collected is not necessarily representative of the general population. Mellon and Prosser (2017) demonstrated that Twitter users differ from the general population in terms of age, gender, and education; the typical Twitter user tends to be young, male, and more educated. While the program utilized in this analysis cannot overcome all concerns about representativeness, these biases can be mitigated through adequate weighting and the sheer numbers of subjects from which data can be collected. Due to the potential sampling pitfalls associated with Twitter as a replacement for traditional survey research, it is important to keep the platform's limitations in mind when planning research. Twitter data may not be suitable as a replacement for traditional research. Instead, Twitter data should be used as a supplement to traditional surveying rather than a replacement for it.

Collecting attitudinal data using tweets, however, also has a clear set of advantages and can effectively address some of the issue raised by other survey approaches, both traditional and online. The collection and analysis of tweets is much faster than fielding and analyzing a traditional telephone-based survey. While a phone survey may take weeks or months to properly execute, this program can collect and analyze thousands of opinions expressed through Twitter within a matter of hours. The speed at which the program can collect tweets is largely dependent on whether the user is searching for tweets in real-time or tweets from a given date. For example, for the purposes of this analysis, I used tweets that were older than the one-week window in which Twitter allows free data collection. I had to purchase the older tweets from Twitter.

Once purchased, the tweets were directly inserted into a database where all of the necessary analyses could be run. Although purchasing tweets added to the cost of the

Twitter analysis, it was still an inexpensive means of surveying when compared with other methods. Rates are dependent on how many tweets are being purchased and the time frame from which the tweets originate. Purchasing less than a million tweets costs around two thousand dollars, which still makes it significantly cheaper than Mturk. Compared to traditional survey methods the dollars-per-subject ratio is the order of magnitudes cheaper.

The most time-intensive part of attitudinal dataset creation using this program is the facial recognition and classification portion of the analysis. The program manually checks and classifies the profile picture of each person in the dataset. Checking each picture and classifying available faces by age, gender, and ethnicity requires a single request for information from the program to the facial recognition and classification application programming interface (API). The program first determines whether there is a suitable face for classification. If the picture does not have a single recognizable face, which is confidently classified by age, gender, and ethnicity, the information is not inserted into the database. It then classifies that face by its respective age, gender, and ethnicity. Finally, it classifies each profile picture by the emotion displayed by the face in the profile picture. This two-step process can take hours. I ran a full image analysis of ten thousand tweets as soon as the initial collection process was completed. It took roughly fourteen hours for the program to examine and classify all of the one thousand tweets. The speed in which this step is completed is dependent upon the speed of the internet connection. The full analysis of ten thousand tweets returned roughly four thousand recognized and classified faces. In short, this program provided me with analyzable data for around 4,000 subjects

Why Twitter?

Perhaps the biggest advantage of using Twitter for research is that it provides access to a rich, largely unfiltered set of data on attitudes. Twitter makes its data easily accessible to its registered developers through its proprietary API. Conversely, Facebook and Instagram have a developer API, but it is closed off to most types of data collection. Many social media platforms use software that prevents web scrapers from compiling data about their users. Twitter's decision to share its data has prompted scholars to devise various means of collecting data from the social media giant.

Other researchers have used Twitter to study a variety of political phenomena, although none of these studies used the tool to the same extent described in this project. Some of the early studies were hampered by low sample sizes, which can probably be attributed to collection methods. Lui, Metaxas, and Mustafaraj (2011), for example, sought to use Twitter as a means of predicting electoral outcomes; however, their study was ultimately undone by its inability to collect a sufficiently representative sample. Evolving technology and a greater understanding of programming and computer science has allowed some researchers to avoid these pitfalls. For example, Tumasjan et al (2010) were able to accurately predict the vote shares of political parties in German elections using data collected from Twitter.

The sheer amount of data available through Twitter has allowed researchers to begin predicting micro-level events as well. For example, Nick Beauchamp was able to use Twitter to predict state-level races with a high degree of accuracy. His 2015 analysis used an impressive collection method combined with sentiment analysis as a means of predicting the electoral outcomes of state-level races. Beauchamp's analysis represents

the most scientific use of Twitter data that has been published up to this point. Many of Beauchamp's techniques are present in this study, but this research innovates upon Beauchamp's work by integrating a more sophisticated duplicate detection method, an improved sentiment analysis, and an image analysis allowing for the collection of more information with a deeper level of sophistication.

As noted above, the key drawback of using Twitter data is that it does not necessarily produce a representative sample. Consequently, there is a reasonable argument that data collected through this method is more of a complementary tool for gathering public opinion data rather than a substitute for traditional survey research. In spite of this, Twitter is still a valuable research tool that should primarily be used for exploratory or supplementary purposes. Researchers can use Twitter to strengthen an argument or to explore social media phenomenon. The benefits of using the program in this manner are significant – if using recent tweets, it is free, and it is fast. Clearly, Twitter can be a valuable tool for researchers who are seeking to better understand public opinion.

Process:

It is important to explain how the web scraping program works on a conceptual basis. The processes used by the scraper can be classified into four major steps: data collection, sentiment analysis, facial analysis, and reading complexity. Understanding these steps is important for verification and replication purposes, but not vital to understanding this analysis or the conclusions reached in this chapter; therefore, these processes are described in greater detail in a technical appendix at the end of the dissertation.

Empirical Analysis:

The program and processes above were used to collect and analyze tweets on Maine's Question Five. The tweets were collected in the month prior to the 2016 election date using the following search terms, which were designed to ensure that I did not miss any useful information. These search terms included:

"Question Five"

"Question 5"

"Ranked-Choice Voting"

"Ranked Choice Voting"

"RCV"

"Preferential Voting"

As noted above, the tweets I wanted to analyze were older than one week, so I had to purchase them from Twitter. In total, I was able to purchase 91,667 tweets containing the above keywords from the month before the ratification of Maine's Question Five. Of these, tweets in which over 50% of the tweet enclosed within quotation marks were excluded from the analysis because these tweets were most likely news-related and not opinions.

The end result was a unique data set consisting of 75,334 unique tweets related to the preferential voting initiative. I used this data to test a set of hypotheses similar to those described in the previous chapter. The dependent variable here is the -1 to +1 score from the sentiment analysis, and the independent variables are age, race, and gender. Kairos estimates age as an exact year estimate, so age was coded as a continuous variable ranging from 16 to 73. Race was coded as a series of dummy variables where white

respondents were coded as one and all nonwhite respondents were coded as zero. These were calculated from the facial analysis as described above. Gender was similarly coded as a dummy variable with one indicating male and zero indicating female. The attitudinal and demographic information available from the data collected allows me to empirically test, with one exception (income), the same hypotheses analyzed in the previous chapter. Specifically:

Hypothesis One: Support for preferential voting will decline with age. In other words, age and emotional polarity will be negatively related.

Hypothesis Two: Minority voters will be more likely to support preferential voting than their white counterparts. In other words, being non-white should positively correlate with emotional polarity.

Hypothesis Three: Female voters will be more likely to support preferential voting than males. In other words, being female should positively correlate with emotional polarity.

Descriptive Results

My initial analysis of the Question Five Twitter data began with an examination of the ratio of positive tweets to negative tweets. I hoped that this comparison would help evaluate whether Twitter is roughly representative of public opinion. The results demonstrated that, once neutral tweets were removed, there were significantly more positive tweets than negative tweets pertaining to preferential voting. Out of the 91,667 tweets, the program found sentiment scores for 75,334 of them. Neutral tweets accounted for 46,777 of those analyzed, which left 31,517 tweets with positive or negative sentiment values assigned by the program.

It should be noted that the number of tweets with accompanying facial data did not vary between sentiment groups. People who tweeted positively about preferential voting were not statistically more likely to post a recognizable display picture than those who tweeted negatively about preferential voting. Likewise, neither group was more likely to post a recognizable display picture than those whose tweets contained no identifiable sentiment. For example, neutral tweets comprised 56% of the entire sample and 59% of the tweets with identifiable faces. Tweets with a positive sentiment comprised 27.7% of the sample and accounted for 25% of the tweets with identifiable faces. Finally, negative tweets comprised 15.75% of the sample and 15% of the tweets with identifiable faces. In other words, the number of identifiable faces is evenly distributed across all sentiment categories.

While the program is equipped with a robust duplicate checking system, this system only applies to the tweets themselves and not to the individuals posting the tweets. Thus, while the program did an adequate job of eliminating duplicate tweets, there were no processes in place to ensure that the program did not collect multiple unique tweets from a single user. As a robustness check, I deleted all identical usernames to see how this would impact the final results of the analysis. Those results will be elaborated upon in greater detail in the next section; however, it is important to note that deleting duplicated eliminated 4,669 observations.

Approximately 62.25% of those remaining tweets were positive, which confirmed the hypothesis that public attitudes as captured by social media posts surrounding the ratification of Maine's Question Five was largely positive, which tracks with the ballot initiative's majority support at the poll. However, the actual voting support for Maine's

Question Five was 52%, considerably lower than the comparable ratio calculated from the sentiment analysis. That suggests the attitudinal picture captured by the analysis of tweets correctly identified the positive direction, but over-estimated support for preferential voting by approximately ten percentage points. While it is difficult to definitively explain this gap, the most likely cause is the potential built-in bias of Twitter's user base, which, as noted above, is not necessarily representative of the general population. These results demonstrate, however, that Twitter can be used to evaluate the directionality of public opinion if the researcher takes the inherent biases into consideration. These results confirm that there was a concerted positive reaction on social media directed towards passing Maine's Question Five.

Hypothesis Tests:

Table 3.1 shows the basic descriptive statistics from the analysis. Table 3.1 (Appendix B) includes columns for the percentages of each demographic found by the program, the raw number of those respondents, the percent of that group who tweeted positively regarding preferential voting. To provide a basis of comparison, U.S. population statistics from the 2010 Census were also included. The program collected and analyzed tweets from a diverse group of individuals. The sample was fairly representative in some but diverged significantly from the U.S. population on some key points. For example, the sample was disproportionately young compared to the census data, which is not hugely surprising given the known demographics of Twitter users (see discussion above).

Table 3.2 (Appendix B) shows the results from a regression analysis similar to the models reported in Chapter Two, though without the income, partisan, employment or

education variables because these could not be extracted from tweet data. The dependent variable in this analysis was whether a respondent tweeted positively about preferential voting. This dependent variable was coded as a dummy variable where One represented a tweet with a positive sentiment and zero represents a tweet with a negative sentiment. The findings presented in Table 3.2 confirm the findings put forth in Chapter Two in substance, but not necessarily insignificant. The results demonstrate that age is a significant negative predictor of a Twitter user's likelihood of tweeting positively about Question Five or preferential voting in general. In other words, older respondents were more likely to tweet negatively. While that supports the finding based on survey data in Chapter Two, the other two variables failed to reach conventional levels of statistical significance. If we interpret coefficient direction, the positive coefficient of the gender variable suggests that males were more supportive of the preferential voting initiative, which is consistent with the finding reported in Table 2.2 but again inconsistent with the hypothesis that females will be more supportive of preferential voting. The race variable is not only statistically insignificant it is in the "wrong" direction and inconsistent with the comparable finding reported in Table 2.2. The positive coefficient suggests whites were more supportive of the preferential voting initiative than minorities. Eliminating the duplicate usernames did not have any significant impact on the results. These results did not vary in any statistically significant way from the results which included the duplicate usernames While these inconsistencies raise obvious questions, the lack of statistical significance cautions about drawing firm inferences from the results.

I suspect that the lack of significance in these results stems from the fact that this is a relatively old sample. As previously mentioned, a tweet is stored with the profile

picture URL present at the time of posting. While I was able to analyze 12,017 faces, this number was cut down significantly by duplicate detection and controlling for quotation marks in tweets. Ultimately, this analysis included approximately 4,602 users (because of missing data on some variables, the total N in the analysis reported in Table 3.2 (Appendix B) is 2,508,). Had I conducted this analysis immediately after the election, it is likely that the program would have had more pictures to analyze. This would have allowed statistically significant results from all explanatory variables.

Discussion:

These results reveal a few key points. First, a significant number of people were tweeting about preferential voting or Maine's Question Five around the time of the election. Additionally, most of these tweets were positive, suggesting majority support for preferential voting, which supports the findings in the previous chapter and, at least in a directional sense, fits with the actual vote on Question Five.

The second analysis in this chapter somewhat confirmed the results found in Chapter Two on specific hypotheses about the demographic traits associated with support for preferential voting. The tweets showed that older respondents were significantly less likely to tweet positively about preferential voting than younger voters, a finding that matches up with the key hypothesis on age and the findings reported using traditional survey data in Chapter Two. The findings on race and gender, however, were mixed. Neither variable was statistically significant. The coefficient for gender suggests males were more supportive of Maine's Question Five, which is consistent with the gender finding reported in Table 2.2 but inconsistent with the operant hypothesis on gender and support for preferential voting. The coefficient for the race variable was positive,

suggesting whites were more supportive of preferential voting, which is inconsistent both with the core hypothesis and the comparable finding from Chapter Two.

As a result of the analysis described in this chapter, I conclude that Twitter can be a useful tool to approximate public opinion data, but there are clearly improvements that need to be achieved before it can be considered an equivalent to traditional survey methods such as those used for the analysis in Chapter Two. The Twitter data accurately predicted majority support for preferential voting, but also clearly overestimated public opinion compared to the actual referendum result. This was almost certainly due to bias in the sample. While the program was unable to yield significant results for every explanatory variable, two of the three variables used in the analysis tracked in the same direction as the comparable variables in the analysis reported in Chapter Two. That suggests that Twitter might be usefully mined to get a reasonable sense of public opinion. Unfortunately, this analysis also demonstrated that extreme bias in Twitter's user population makes it difficult to use the social media platform to draw any substantive conclusions regarding public opinion. Instead, this analysis demonstrates that Twitter is best used for supplemental or experimental analyses.

Chapter Four:

The previous two chapters provide evidence that a majority of voters support preferential voting, and that those levels of support may vary on the basis of socio-demographic traits. Specifically, groups that have been traditionally marginalized and less likely to benefit from the political outbidding endemic to plurality systems are typically more likely to favor a shift to a plurality system. While the empirical evidence was somewhat mixed when it came to individual variables, overall the general pattern was of higher levels of support for plurality voting among these groups.

Such levels of support, however, do not automatically mean plurality voting will address the concerns these groups have about plurality systems. Political outbidding typically translates into a sort of "us versus them" style of political campaigning, one where candidates make promises and pledges to certain groups that often come at the expense of others. While such promises can be aimed at traditionally marginalized voters (e.g. affirmative action), they are more likely to be directed at dominant voting blocs for the simple reason that is where the most votes are. If preferential systems do indeed address these sorts of issues, then we should see a lot more than just public support for such systems, and systematic variation across groups in levels of support. Preferential systems should actually change specific electoral incentives for candidates, and in doing so change campaign styles and tactics in predictable ways. The primary objective of this chapter is to examine exactly this issue, i.e. the impact that the transition to preferential voting has on the campaign process.

Specifically, this chapter will focus on the differences between preferential cities and plurality cities regarding campaign negativity and campaign tactics and the

theoretical reasons why those differences exist. This portion of the project will explore these topics through the continued utilization of the Eagleton poll data (i.e. the same data sources used in Chapter 2).

Campaign Effectiveness:

Candidates in both plurality and preferential voting systems have strong incentives to engage with potential supporters in the most effective way possible. In this context, campaign effectiveness is defined as the impact of voter mobilization efforts. If they seek to win an election, candidates have an obvious incentive to engage in actions that they believe will increase the probability of getting people to actually show up and cast a vote supporting their candidacy. More effective campaigns are those that engage in behaviors that best translate into support at the polls

One of the seminal volumes on the effectiveness of campaign tactics is Gerber and Green's (2014) *Get Out the Vote: How to Increase Voter Turnout*. In their book, Gerber and Green describe their experiments which test the effectiveness of various campaign strategies in a real-world election setting. Gerber and Green measure campaign effectiveness in bang-for-buck terms (number of dollars spent per vote). When viewed through this lens, some tactics, such as door-to-door canvassing, are extremely effective despite the high upfront costs, because they produce the most votes. Other strategies, such as online advertising, may cost less than door-to-door canvassing but are less effective in that they result in fewer votes. Although most of the experiments described in *Get Out the Vote* took place during a presidential election and the data used in this project focuses on local elections, Gerber and Green's research provides valuable information regarding the effectiveness of various campaign tactics in a real-world setting.

There is a wide range of tactics that can be employed to try and achieve the goal of campaign effectiveness as described by Gerber and Green. Door-to-door canvassing, or interacting with constituents in person, is generally acknowledged as the most effective method for increasing voter turnout (Gerber et al 2003, Gerber & Green 2014, Michelson 2003). Gerber and Green (2014) estimated that door-to-door canvassing produces one vote for every fourteen contacts and costs the candidate $29 for every vote. In contrast, phone banking, another commonly employed campaign activity, produces one vote for every 38 contacts and costs the candidate $38 for every vote. Gerber and Green report that other forms of campaigning, such as e-mail, television, and radio are frequently used to disseminate a candidate's message, but these methods have not been shown to have a statistically reliable impact on a candidate's electoral performance.

Negative Campaigning:

It is important to note that campaign tactics and campaign tone are not mutually independent. Television advertisements may not be statistically proven to impact a candidate's electoral performance, but negative TV ads are a standard campaign activity. Voters commonly associate attack advertising with television advertisements (Fridkin & Kenny 2004). However, campaign negativity may be found in many other campaign strategies. Push polls, for example, are a campaign tool in which a seemingly innocuous telephone survey is used to disguise a negative campaign message (Randolph 2008). Similarly, as demonstrated during the 2016 election, seemingly innocuous social media groups can be a front for negative campaign tactics and attack advertising (Wong 2017) Campaign negativity is deeply embedded into the election experience in the United States. Nearly every presidential election is described as being the most negative election

in history. Of course, intrinsic in this observation is the understanding that the following election will most assuredly be perceived as more negative than the one that came before it. In July 2016, Aaron Blake, a reporter for The Washington Post, wrote: “It’s become cliché to decry each election as the most negative of our lives . . .” (Blake 2016). Blake then stated, “Polling shows voters indeed are already more prepared to vote against something than for something in 2016. A Pew poll from February 2018 showed 50 percent of Clinton’s supporters said their vote was mostly against Trump, while 55 percent of Trump supporters said their vote was mostly against Clinton.”

The scholarship is conflicted regarding the effectiveness of negative campaigning in plurality voting contests. Some scholars contend that negative campaigning is an effective campaign strategy for mobilizing a candidate’s base or at least demobilizingthe base of the candidate’s opposition. Ted Brader (2005) effectively demonstrated that fear can stimulate voter vigilance and thus influence voter behavior. Other scholars have asserted that the relationship between voter demobilization and campaign negativity is dependent on which candidate the negativity is directed towards and which candidate is preferred by the voter (Krupnikov 2011).

Conversely, numerous scholars contend that negative campaigning does little to motivate voters; these scholars assert that negative campaigning tends to suppress voter turnout and decrease the overall mood of the electorate. Voters who perceive a campaign as being extremely negative are more likely to report decreased feelings of political efficacy and reduced trust in government (e.g. Lau et. Al. 2007).

While the existing research makes it difficult to assess the exact impact of negative campaigning on voter turnout, there is no disagreement about its impact on

public attitudes. A wealth of scholarship demonstrates that excessive negative campaigning is not viewed favorably by the public. Independent voters are likely to shy away from engaging in campaigns they perceive as being highly negative (Lau and Pomper 2004). Attack advertising can have detrimental impacts aside from demobilizing key voting demographics. Negative campaigning can reduce voters' willingness to seek out information pertaining to elections (Shah et al. 2007). While the effectiveness of negative campaigning in terms of increasing/decreasing vote share and turnout is contested, the unpopularity of negative campaigning is well-documented and almost universally accepted.

Campaigns are historically inventive when it comes to attack advertising. Preferential voting ultimately decreases the utility of these campaign tactics by encouraging a larger field of viable candidates—even in a system with two dominant political parties, preferential systems provide clear incentives for greater participation by third-party and/or independent candidates because it reduces the need for strategic voting. In other words, those minor parties and candidates can more successfully seek votes in a preferential versus a plurality system. A greater pool of candidates in a plurality system has implications for the attack advertising campaign efforts so common in contemporary electoral contests. After all, it is much easier to rely on attack advertising when a candidate only has one or two viable opponents; however, attack advertising becomes less useful when a candidate has, say, five or six viable opponents. As the number of viable candidates increases the viability of attack advertising decreases and the utility of voter mobilization increases.

In a plurality system, elections are often a contest between two opposing candidates, both of whom have clear incentives to employ negative campaign appeals (see discussion below). Increasing partisan polarization increases the emphasis on negative messaging beyond the context of any given election. This polarization is not only limited to elections. Party polarization has reached a 100-year high in Congress (Ansolabehere & Iyengar 1996), and the increasingly "us vs. them" nature of politics in the United States leaves little room for moderates and centrists (Abromowitz & Webster 2015).

Candidates in cities with a preferential voting system find themselves in a very different political environment, one that provides incentives for less oppositional campaign strategies. Preferential voting systems may decrease the potential payoffs of negative campaigning because of the offer fewer incentives to engage in negative campaigning than do plurality systems. Preferential voting is not a zero-sum game with only one winner. It can be advantageous for a candidate to be a voter's second or even third choice, so negative campaigning becomes less valuable due to the increased number of candidates and the increased number of candidates for which a citizen may vote. Preferential voting encourages reciprocity amongst rival candidates who all share a mutual desire to be selected by a potential voter; it thus makes sense that candidates work together instead of against each other (Norris 2004; Neely et al. 2005). Preferential elections may similarly decrease attack advertising by encouraging candidates to dedicate more resources towards mobilization (Horowitz 1985). The increased number of candidates usually present in a preferential election lessens the value of attack

advertising. As the number of candidates increases, it becomes more difficult to attack all the candidates (Donovan 2003).

Researchers have provided empirical evidence proving that preferential voting contests are less negative than plurality campaigns. For example, Todd Donovan (2003) demonstrated that ranked choice voting dramatically reduced perceptions of campaign negativity while improving perceptions of campaign civility within the general electorate. Donovan surveyed candidates in preferential voting elections and asked them about their perceptions regarding campaign negativity. Donovan's findings were confirmed in subsequent research in conjunction with Caroline Tolbert and Kellen Gracey (2016). Tolbert, Gracey, and Donovan (2016) found that voters in preferential elections were less likely to indicate that they had heard candidates criticizing each other in local elections. Tolbert, Gracey, and Donovan (2016) tested hypotheses similar to the hypotheses I am testing in this Chapter; however, it should be noted that their examination of negativity in preferential voting relied on an older dataset, different models, and did not examine the relationship between preferential voting and campaign tactics.

Hypotheses:

Based on the discussion above, this chapter will test a series of hypotheses positing that, compared to plurality systems, electoral campaigns in preferential systems will be less negative and encourage more effective campaign behaviors by investigating voters' perceptions of candidate behavior and the campaign tactics used in the cities contained within the dataset. The specific hypotheses to be tested are as follows:

Hypothesis One: Voters in cities with preferential voting will report lower perceptions of campaign negativity than voters in cities with plurality voting.

Hypothesis Two: Voters in cities with preferential voting are less likely to report instances of candidates criticizing each other than voters in cities with plurality voting. Voters in cities with preferential voting are also expected to be more likely to report instances of candidates praising each other than voters in cities with plurality voting.

Hypothesis Three: Voters in cities with preferential voting are more likely to report being contacted in person by someone representing a candidate than voters in cities with plurality voting.

Hypothesis Four: Voters in cities with preferential voting are more likely to report that they were contacted by campaigns (either directly or indirectly) more often than voters in cities with plurality voting.

The first two hypotheses are largely focused on how voters perceive the tone of a campaign, and the idea is to investigate whether voters in the preferential system see campaigns as more positive (or at least less negative) than campaigns in plurality systems. Hypotheses three and four focus less on the relative negativity of elections and more on the tactics undertaken by candidates' campaigns. These latter two hypotheses are squarely aimed at the question of whether plurality systems are more likely to encourage effective campaign tactics as defined by Gerber and Green (see discussion above). If the empirical evidence supports these hypotheses it would suggest that plurality systems are less negative and incentivize campaign behaviors that maximize political participation. If this is indeed the case it would not only suggest that the support for such systems by

marginalized voters (see previous two chapters) is well-founded, more broadly it would imply preferential systems increase participation, reduce negativity and address a number of the concerns raised about the impact of plurality systems on trust and efficacy discussed in earlier chapters.

Research Design:

Campaign Tactics:

To test the hypotheses relating to campaign tactics I use a set of survey questions probing if and how candidate campaigns had connected with voters. The survey asked respondents to state whether or not they had been contacted by a campaign. Respondents were asked whether or not they had been contacted through a variety of campaign mediums, such as television advertisements, email, mailers, in-person contact, or phone calls from campaigns. The specific questions are as follows:

"Next I will read a list of ways a campaign or candidate might have contacted you. For each way simply tell me yes or no. Was the contact:

> *By telephone?*
>
> *By printed mail to your home address?*
>
> *In person, either at your house, or in public? By email?*
>
> *Through a social network site like Facebook or Twitter?"*

I used these to create two dependent variables relevant to Hypothesis Three and Hypothesis Four. The first variable was a simple dummy where respondents who stated that they were contacted in person were coded as one, while respondents who stated that they were not contacted in person were coded as zero (Hypothesis Three). This coding

strategy was also designed to distinguish between respondents who were contacted by the most effective campaign tactic and those who were not.

The second variable was used to operationalize the key concept in Hypothesis Four, i.e. the degree of contact that voters receive from campaigns. Accordingly, I created a continuous variable by combining all of the binary variables pertaining to respondent's interactions with specific campaign methods. This created a continuous variable ranging from 0, which represents respondents who did not interact with any campaign method, to 6, which represents respondents who were contacted through all campaign mediums.

This model utilized ANOVA and a subsequent Tukey's HSD test as a means of examining differences between preferential cities and plurality cities. The decision to use an ANOVA was made because a respondent's socio-demographic status could influence how they were contacted by a campaign. Plus, there are a variety of mediums which may not be readily available to large swaths of the population. For example, an increasingly large number of voters do not have cable television; conversely, many voters do not have a landline telephone. There were too many confounding variables that I could not control for; therefore, I chose to use an ANOVA because it would allow me to compare two group means. In this case, the ANOVA is comparing the group means of cities with preferential voting and cities with plurality voting. While the survey contained a number of socio-demographic measures, there are numerous factors which could influence a respondent's ability to be contacted by a campaign which here not measured by the survey. For example, the survey did not ask respondents whether they have any form of social media, whether they have an email address, whether they have a landline, or whether they live in an apartment or house. All of these factors could influence themeans

by which a candidate was able to contact a respondent; therefore, an ANOVA was suited as a means of broadly comparing campaign strategies between cities, as it would be impossible to control for confounding variables adequately.

Campaign Negativity:

To investigate the first two hypotheses described above I use data from the Eagleton Poll already described in Chapter Two. The data comes from a 2014 survey of cities with preferential voting and cities with plurality voting in California. The survey had 2,456 total respondents (see chapter two for a more in-depth description of this data). This survey contained questions that permit fairly straightforward tests of the first two hypotheses. Specifically, respondents in all cities were asked:

"*Do you believe the campaigns this year were more negative, less negative, or about the same compared to other recent political contests?"*

Respondents were also asked two follow-up questions asking for clarification of whether the campaigns were "a little" or a "lot more" negative or positive. Responses were combined to create a 5-point ordinal scale that ranged from campaigns "were a lot more negative" (coded 1), "a little more negative" (coded 2), "about the same" (coded 3), "a little less negative" (coded 4) and "a lot less negative" (coded 5). This combined ordinal variable was used as a measure of the dependent variable for an analysis of perceived negative campaigning, i.e. to test Hypothesis One. The same survey also asked the following questions:

"*During the recent election do you remember examples of candidates praising or endorsing any of their opponents?*"

"*Thinking about the recent election, how much time would you say the candidates spent criticizing their opponent?"*

These were used to create the key dependent variable to test **Hypothesis 2**. Responses to the first questions were coded ordinally from "Yes, frequently" (coded 4) to "No, never" (coded 1). Responses to the second question were coded from "A great deal of the time" (coded 4) to "They weren't doing this at all" (coded 1). This creates an ordinal measure of negativity where respondents could signify whether they perceived increased amounts of negativity to decreased amounts of negativity.

The key independent variable used to explain variation in the dependent variables described above is a dummy variable signifying whether the city a respondent is from uses preferential or plurality voting. The general expectation is that respondents in cities with preferential systems will perceive campaigns as less negative, will be more likely to report hearing candidates praising each other, and less likely to hear candidates criticize each other.

The statistical models used to test the key hypotheses also included a series of control variables. These included age, gender, ethnicity, and partisanship. The analysis controlled for income, which included dummy variables for low-income respondents, high-income respondents, and a variable for respondents who refused to specify their income level. This was done as a means of controlling for respondents who refused to specify their incomes. This approach allows me to include respondents who did not include their income without imputing their respondents and potentially damaging the accuracy of the analysis. This strategy allows me to use all available information about a missing observation. This strategy has been utilized by multiple scholars and is an

accepted means of dealing with missing data (Bhat 1994). The study also controlled for education, employment status, and marital status. Employment and marital status were included as control variables because both can have a measurable impact on an individual's ability to become politically involved (Rohe & Stegman 1994). Since being politically involved would be a prerequisite for observing candidates engaging in the measured behaviors, these variables were included. The use of a multivariate model allows me to control for factors to ensure that any disparity between cities with preferential voting and cities without preferential voting is due to preferential voting and not other contextual factors. This analysis is primarily concerned with understanding whether preferential voting has an impact on people's perceptions of campaign negativity and candidate behavior. Therefore, the models designed to examine this question were designed with an ordinal dependent variable which asked respondents to rate their perceptions of campaign negativity in their city.

Findings

Results from the models testing the first two hypotheses (those dealing with campaign tone or negativity) are reported in Tables 4.1, 4.2 and 4.3 (Appendix B). Table 4.1 shows the results from an ordered logistic regression which indicate that respondents in cities with preferential voting report lower perceived levels of campaign negativity than respondents in cities with plurality voting. This finding of lower negativity in preferential systems held across all models reported in Table 4.1. Additionally, the results showed that younger voters and more educated voters perceived less campaign negativity than individuals from other demographics. Predicted probabilities demonstrate that respondents in cities with preferential voting have a 4% greater chance of reporting "a lot

less negativity" than their counterparts in cities with plurality voting. Conversely, respondents in cities with plurality voting had a 4% greater chance of reporting "a lot more negativity" than respondents in preferential cities. These findings largely confirmed my hypothesis. Cities with preferential voting exhibit lower levels of campaign negativity than cities with traditional plurality voting.

As shown in Table 4.2, there was no significant difference between cities with preferential voting and cities with plurality voting regarding candidates criticizing each other. In other words, voters did not perceive a significant difference in how much candidates criticized each other regardless of which voting system their city was using. White respondents were less likely to report hearing candidates criticize each other than nonwhite respondents. Employed respondents were more likely to report hearing candidates criticizing each other, which may be a spurious correlation or the result of employed respondents having more money, which usually correlates the greater political involvement. Table 4.3 shows that candidates in cities with preferential voting were more likely to be heard praising each other. White respondents were less likely than nonwhite respondents to report that they had witnessed positive campaign strategies. These regressions mirror the findings put forth by the primary analysis — preferential elections are less negative than plurality elections.

The results of the analyses testing hypotheses three and four (dealing with campaign tactics) are reported in Tables 4.4 and 4.5. Table 4.4 indicates that cities with preferential voting are significantly more likely to employ campaign methods that focus on in-person methods of contact. These results were significant at the .05 level.

Conversely, citizens in cities with plurality systems did not enjoy this same level of person-to-person contact with the candidates. Table 4.5 demonstrates that respondents in cities with plurality voting were statistically less likely to report having engaged with a campaign in any capacity. This confirms the hypothesis that candidates in cities with preferential voting are more likely to focus on campaign methods that have been shown to effectively mobilize voters. The results indicate that plurality cities had a mean of 46.44%, while cities with preferential voting had a mean of 51.54%. This means that respondents in cities with preferential voting were roughly 5.1% more likely to report being contacted by a campaign in person.

These analyses confirm my initial set of hypotheses. Respondents in cities with preferential voting are more likely to report lower levels of perceived negativity than respondents in cities without preferential voting. The results indicate that there are significant and measurable differences in campaign tactics used in cities with plurality voting and cities with preferential voting. These results indicate that the styles of campaigning generally deemed effective in winner-take-all political systems are not necessarily as effective in cities with preferential voting. The data indicates that respondents in cities with preferential voting are significantly more satisfied with the campaigns being run in their respective cities.

Conclusions:

The analyses in this chapter demonstrate that there are significant differences between the elections conducted in cities with preferential voting and those in cities that use plurality voting. This analysis focused on the differences between these two types of elections during the campaign process. The analysis successfully demonstrated that there

are differences between the two types of elections regarding campaign negativity and campaign tactics. Respondents in cities with preferential voting are statistically likely to report lower levels of perceived campaign negativity.

These findings indicate that candidates in cities with preferential voting must contend with different political pressures than candidates in plurality contests. As the number of candidates in an election increases, it becomes increasingly difficult to use negative campaigning effectively. Instead, candidates in cities with preferential voting likely rely on campaign strategies that seek to mobilize their voters rather than those that seek to demobilize their opponent's voting base. Second, given the nature of preferential voting, it is beneficial for a candidate to avoid upsetting his/her opponent's voting base since being a voter's second or third choice is still valuable for a candidate.

The analysis also demonstrated that there are significant differences in campaign strategies used in cities with preferential voting contests and those with plurality voting. Candidates in cities with preferential voting were more likely to focus on more effective campaign tactic, which is often associated with voter mobilization rather than attack advertising. Preferential voting candidates are more likely to use campaign tactics that involve in-person contact. Additionally, respondents in cities with preferential voting were more likely to indicate that they have been contacted by a campaign through more mediums than respondents in cities with plurality voting. The results indicate that the systemic differences between preferential systems and plurality systems require and encourage different campaign strategies to help candidates accrue as many votes as possible.

Chapter Five:

The key takeaways from the previous three chapters are that, (a) there is strong support for preferential systems among voters that is detectable using traditional survey data and data collected from social media; (b) that support systematically varies across particular socio-demographic groups in patterns consistent with the theoretical expectations laid down in Chapter One and Chapter Two, though this picture that is much clearer using traditional survey data as opposed to the data scraped from social media; and (c) that preferential systems seem to promote less negative and more effective campaigns, elections with higher rates of voter engagement and participation and more positive in tone that plurality voting systems. Normatively speaking, the findings thus paint a fairly positive picture in favor of preferential voting, at least in the in the sense of boosting civic engagement in the democratic process, reducing political outbidding, and generally promoting a less negative tone.

None of the findings thus far, however, address one of the key criticisms of preferential systems: the added complexity and information burden they place on voters compared to plurality systems. In this chapter, I seek to address two fundamental questions pertaining to this issue. First, can citizens understand the basic concept of preferential voting and how it works, at least to the same degree that they understand plurality voting? Second, what explains any variation in such understanding—are some groups of voters more likely to understand plurality voting than others?

The first question is important to address because the most frequently-cited criticism pertaining to preferential voting is that the American electorate simply lacks the political sophistication to engage with any system of voting other than plurality voting.

These critics posit that allowing a voter to cast a vote for more than one candidate requires a level of political knowledge that most Americans simply do not possess. For example, even though it was approved by voters, there was considerable opposition to the Maine plurality voting initiative, and that opposition was driven at least in part by questions about whether voters could fully grasp the mechanics of plurality voting. For example, Gordon L. Weil, a writer for the *Portland* (Maine) *Press Herald*, referred to preferential voting as "Costly, Complicated, and Undemocratic" (Weil 2016). Weil contended that, among other things, preferential voting is too complicated for the average voter. He asserted that many voters lack the sophistication to engage in preferential voting. Weil's concerns are not uncommon. Many believe that the majority of Americans are hardly capable of choosing one candidate to vote for, let alone three candidates.

Empirical support for such criticism, however, is limited. There have been multiple studies exploring voters' understanding of plurality voting (e.g. Lee et al. 2004, Blias et al. 2006), and most of these studies have demonstrated that a voter's ability to understand plurality voting is heavily dependent his or her level of education. More educated respondents are more likely to understand plurality voting (Lee et al. 2004). Conversely, relatively little research has been done to examine whether understand preferential voting. This lack of academic research is largely due to a dearth of solid data and the relatively small number of cities currently using preferential voting. There is no extant research which attempts to demonstrate that voters do understand preferential voting in the cities where it has been implemented. Neither have any researchers attempted to determine if a voter's ability to understand preferential voting differs

significantly from his or her ability to understand plurality voting. Therefore, this chapter represents a significant and useful entry into the existing body of literature.

The second question in this chapter pertains to who understands preferential voting. Again, the assertion of some critics is that voters' understanding will be adversely impacted by the process of rank ordering up to three candidates as opposed to selecting a single candidate (Neely et al. 2005; Cook & Latterman 2011; Arrow & Raynaud 1986).

Such criticisms come not just from academics, but also from political elites. Jerry Brown Jr, Former Mayor in Oakland and the governor of California said: "In a time when we want to encourage voter participation, we need to keep voting simple. Ranked choice voting is overly complicated and confusing. I believe it deprives voters of a genuinely informed choice." Katie Blinn, assistant state director of elections in California, stated: "People didn't understand how the votes were counted. These examples are representative of the numerous critics of preferential voting who assert that preferential voting is too complicated." Whatever the merits of these criticisms, it is almost certainly the case that not all voters are equally informed. What might explain individual-level variation in how preferential voting works?

Hypotheses:

While many critics decry preferential voting as being too complicated for voters, my first hypothesis is that there will be no statistical differences between the groups who understand preferential voting and the groups who understand plurality voting. This hypothesis is anchored in the assumption that certain socio-demographic groups have little understanding of voting systems generally, and that this inability to understand is

not dependent on the type of voting system with which they are required to engage, but more on the information incentives embedded in those systems. In a typical plurality system where outbidding is common, typically marginalized voters have low incentives to bear the information costs to fully understand the mechanics of that system. Such systems, as argued earlier in this dissertation, tend to be less attentive to their preferences because of how plurality voting structures incentives for campaign behavior by candidates. Plurality systems, on the other hand, are more likely to attend to their preferences and thus may alter the information incentives for these marginalized voters. This assumption is backed by contemporary scholarship, which demonstrates that political participation and interest can increase among politically disenfranchised minorities when those minorities perceive themselves as having an increased stake in the political system (Just 2017). Thus, politically disenfranchised minorities will learn about and master the preferential voting process if they believe it can be used to increase their political capital. If this is so, at a minimum, voters are unlikely to understand preferential systems any less than they understand plurality systems.

While this chapter's first hypothesis examines respondents' ability to understand the concept of preferential voting, Hypothesis Two examines respondents' self-reported ability to understand the specific instructions that they were required to read prior to voting. For reasons just discussed, I predict that typically marginalized voting groups will be able to understand these instructions at least as well as voting groups that typically represent majorities. Specifically, I will test whether ethnicity, gender, or income significantly predicts a respondent's ability to understand these voting instructions. I chose these socio-demographic characteristics because they are most often

associated with political disenfranchisement. This hypothesis is anchored in the assumption that politically disenfranchised minorities have incentives to educate themselves sufficiently to understand the instructions required of them prior to engaging in the preferential voting process. This means that the respondents captured by this survey question have likely already educated themselves about their respective voting system. This hypothesis is meant to test whether these respondents were not confused by their voting system, but more specifically by the instructions, they were asked to read.

The analysis utilizes the same 2014 FairVote dataset referenced in the previous chapters. This data was gathered by the Eagleton Poll. The hypotheses tested in this analysis are both included in a single study. Based on the above discussion, the key hypotheses to be tested are as follows:

Hypothesis One: There will be no significant differences between socio-demographic groups regarding their ability to understand preferential voting and plurality voting. Specifically, gender, income, and ethnicity will not predict an understanding of preferential voting.

Hypothesis Two: The socio-demographic characteristics typically associated with political disenfranchisement (gender, ethnicity, and income) will not be significant predictors of a respondent's ability to understand voting instructions

Analysis:

I tested Hypothesis One using three separate statistical models. Two statistical models will be focused on discerning which socio-demographic characteristics predict understanding the concepts and tenets behind preferential voting and plurality voting.

Additionally, these models will test to see if the respondents who purport to understand preferential voting are different than the types of respondents who claim to understand plurality voting. These two models will use two ordinal variables designed to test respondents' understanding of preferential voting and plurality voting. Two questions were used as the dependent variables for these two models. Respondents were asked:

"Overall, how well do you think you understand ranked-choice voting?" and *"Overall, how well do you think you understand plurality voting?"*

Respondents were asked to rank their understanding of preferential and plurality voting on an ordinal scale ranging from "not at all" (coded as One) to "extremely well" (coded as Four) It should be noted that the question respondents were asked was dependent on the type of voting present in their city. Respondents in cities with preferential voting were asked only about their understanding of preferential voting, and respondents in cities with plurality voting were only asked about their understanding of plurality voting. This question is somewhat flawed in that social desirability may play a strong role in how respondents answered this question. Obviously, nobody wants to admit that they hardly understand the voting system in which they are about to participate; however, despite this, answers were normally distributed with nearly half of respondents (47.78%) admitting they understand preferential voting "not at all well" or "somewhat well."

These two models used ordinal logistic regressions and a variety of demographic indicators to isolate the characteristics associated with voters who can or cannot understand their city's respective voting system. The models control for partisanship, a respondent's level of education, marital status, and employment status. Partisanship was

included because, as demonstrated in Chapter Two, partisanship can have an impact on a respondent's receptiveness to preferential voting. While I doubt partisanship will have a significant impact on this model, I believe that it is an important variable to control for given the previously demonstrated relationship between partisanship and alternative voting systems. Education is included as a control variable because more educated voters are expected to have a greater understanding of whatever voting system they participate in. A dummy variable for "missing income" was included to control for respondents who chose not to share their incomes. This "missing income" variable was included as a means of including respondents who did not provide their income but still provided information valuable to the analysis. Given that one of the goals of these models is to determine whether minority voters are disproportionately impacted by preferentialvoting, the models coded ethnicity as a dummy variable where white respondents are coded as one and all non-white respondents are coded as zero.

The results from the analysis on how well voters understand voting systems are displayed in the first two columns of Table 5.1 (see Appendix B). The results indicate that there are certain groups which have a difficult time understanding preferential voting and that demographic characteristics do not necessarily predict difficulty in understanding plurality voting. These results undermine support for Hypothesis One. First, high-income respondents, well-educated respondents, and male respondents were significantly more likely to understand preferential voting. While high-income respondents were statistically more likely to understand preferential voting, low-income respondents were not statistically less likely to understand preferential voting. In other words, having higher levels of income can correlate to a greater understanding of

preferential voting, but less education does not correlate with an inability to understand preferential voting.

As a supplemental analysis, I tested the relationship between a respondent's understanding of preferential voting and his or her support for the implementation, or continued implementation, of preferential voting in their city. When I used a respondent's understanding of preferential voting as an explanatory independent variable in a model identical to the logistic regression utilized in Chapter Two, the results demonstrate that understanding preferential voting is a significant predictor of someone's willingness to support it. Respondents who understand preferential voting "not at all well" had a 22.9% chance of supporting preferential voting, while respondents who understood preferential voting "extremely well" had a 55.6% chance of supporting preferential voting.

The results regarding plurality voting were similar to those for preferential voting. Respondents with more education were more likely to understand plurality voting. As shown in Table 5.1, older respondents were more likely to understand plurality voting, and unemployed respondents were less likely to understand plurality voting.

Table 5.2 summarizes the results from the regression models for understanding preferential voting. The results demonstrate that high-income respondents were statistically more likely to understand preferential voting. Respondents with higher levels of education were more likely to report that they understood preferential voting. Finally, male respondents and employed respondents were more likely to report that they understood preferential voting. These results demonstrate that, unfortunately, there

appears to be a significant economic component tied to a respondent's ability to understand his/her respective voting system. Furthermore, the fact that men are more likely to understand preferential voting than their female counterparts demonstrates that, despite its popularity with traditionally disenfranchised groups, preferential voting is not a universally understood concept.

Table 5.2

More Likely to Understand Preferential Voting	**Less Likely to Understand Preferential Voting**
High-Income Respondents (Dummy Variable – High-income was significant, but low-income was not)	-
More Educated Respondents (Ordinal – Likelihood of understanding increases with more education)	-
Male Respondents	Female Respondents
Employed Respondents	Unemployed Respondents

Table 5.3 summarizes the regression model results for understanding plurality voting. Similar to preferential voting, education is a significant predictor of a

respondent's likelihood to understand plurality voting. Older respondents were more likely to understand plurality voting, which makes sense given that older respondents have had more exposure to plurality voting than their younger counterparts. Again, similar to preferential voting, employed respondents were more likely to report that they could understand plurality voting. Finally, married respondents were more likely to understand plurality voting than unmarried respondents. This could be related to the fact that married respondents tend to be older or have higher income levels. It should also be noted that the dummy variable controlling for missing income was statistically significant; however, it is difficult to interpret this finding given that the variable is literally controlling for a lack of data. Given that higher income respondents are often less likely to report their income on a survey (Turrell 2000), it would seem likely that this is suggesting that higher income respondents were more likely to understand plurality voting.

Table 5.3

More Likely to Understand Plurality Voting	**Less Likely to Understand Plurality Voting**
More Educated Respondents (Ordinal – Likelihood of understanding increases with more education)	-

Older Respondents (Continuous - Likelihood of understanding increases as the age of the respondent increases)	-
Married Respondents	Unmarried Respondents
Employed Respondents	Unemployed Respondents

The results shown in tables 5.2 and 5.3 indicate that education and income are the key traits predicting an individual's ability to understand his/her respective voting system. Unemployed respondents consistently express less understanding of voting systems, regardless of in which system they participate (or do not participate), and somewhat surprisingly, there is a gender difference in understanding preferential (but not plurality) systems. Ultimately these findings illustrate the impact of socio-demographic factors as it pertains to facilitating an understanding of politics and increasing political efficacy.

The second analysis focused on whether voters in cities with preferential voting were able to understand the instructions they were asked to read when voting. The results from this analysis are found in the third column Table 5.1 and are summarized in Table 5.4. This analysis uses an ordinal logistic regression. The dependent variable in this analysis asked respondents to respond to the question:

"When you voted in the recent election, how easy was it to understand the voting instructions?"

Respondents were asked to rate the ease of understanding voting instructions on a four-point scale ranging from “very Difficult” (Coded One) to “very Easy” (Coded Four). Again, this model controlled for age, ethnicity, gender, partisanship, income, missing income, and a respondent’s level of education. The sample will be subset between cities with preferential voting and cities without preferential voting as a means of focusing on specific voting systems.

Table 5.4

More Likely to Understand Instructions for Preferential Voting	**Less Likely to Understand Instructions for Preferential Voting**
White Respondents	Minority Respondents
Male Respondents	Female Respondents

Conclusions:

This chapter has demonstrated that some groups do indeed have more difficulty understanding and engaging with preferential voting. Male respondents were more likely to indicate that they understood preferential voting, which means that, in this context of this analysis, female respondents were statistically less likely to understand preferential voting. While I did not account for gender in my theory section, it is worth considering why female respondents were less likely to understand -- or less likely to admit that they understand -- preferential voting. I suspect that female respondents did not, in fact, differ

from their male counterparts with regard to their actual ability to understand preferential voting. I do suspect that female respondents varied from male respondents in their willingness to admit that they do not understand preferential voting. Women, in general, are more likely to admit that they don't understand something, while men are more likely to say that they understand something when they don't (Babcock & Laschever 2009). This gendered social desirability effect is unfortunate but could explain the discrepancy between male and female respondents in their reported levels of understanding.

Respondents with more education and more income were also more likely to understand preferential voting. Similarly, male respondents and white were more likely to understand the instructions they were required to read prior to voting in a preferential election. These findings indicate that there are socio-demographic groups that are disadvantaged by preferential voting and have difficulty understanding it.

Similar to preferential voting, there are socio-demographic characteristics which predict a respondent's ability to understand plurality voting. Certainly, more educated respondents are more likely to understand it. These results mirror some of the findings which indicate that education benefits people's ability to understand voting systems generally. Older respondents were more likely to understand plurality voting, which makes sense given that older respondents have been exposed to plurality voting longer than younger respondents. Education and employment status were predictive of respondents' ability to understand preferential and plurality voting; however, income and gender were significant predictors in the preferential voting analysis but not the plurality voting analysis. Conversely, age and marital status were significant predictors in the plurality voting analysis, but not the preferential voting analysis.

Contrary to expectations, at least some of the criticisms directed toward preferential voting have been empirically validated by this chapter. Specifically, the analysis has demonstrated that female respondents and minority respondents are more likely to have difficulty understanding preferential voting and/or the instructions they are required to read prior to voting in a preferential election. Respondents with lower levels of education were less likely to understand preferential voting; however, the results demonstrate that respondents with lower levels of education were less likely to understand plurality voting, so this is not unique to a single type of voting system. There are socio-demographic characteristics who will be disadvantaged regardless of which voting system a city or state chooses to adopt; therefore, it falls on the city or state to examine their population and decide if the benefits outweigh the potential costs.

Chapter Six:

This dissertation has focused on preferential voting in the United States. The goal of this analysis was to investigate how voters perceived preferential voting, to explore how the transition from plurality voting to preferential voting impacts the campaign process, and to assess whether voters could bear the information costs associated with understanding how a preferential voting system works compared to a plurality system. This chapter serves as the conclusion to the rest of the analysis and will summarize the findings presented in this dissertation. This chapter will conclude with a brief exploration of the potential future of preferential voting in the United States.

Chapter Two isolated the socio-demographic characteristics associated with those who support preferential voting. Additionally, this chapter explored whether voters in cities with preferential voting supported the continued implementation of preferential voting, and the chapter demonstrated that respondents support its continuation. This finding demonstrates that voters appreciate having the opportunity to rank order their candidates. This chapter also demonstrated that the socio-demographic characteristics typically associated with political disenfranchisement and political outbidding (ethnicity, age, and income) tend to predict support for preferential voting. While this support could be due to many factors, these findings are clearly consistent with the theoretical framework laid down in Chapter One and Chapter Two, especially in regards to political outbidding and marginalized voters. Preferential voting reduces or eliminates the need for political outbidding, which favors voters in the majority, and gives traditionally marginalized voting blocs more incentive to engage and participate. Preferential voting does not force voters to choose between a candidate who represents their ethnicity and a

candidate who represents their policy preferences; instead, voters in preferential systems can choose on the same ballot candidates who represent their ethnicity as well as candidates who represent their policy preferences. Finally, younger voters may prefer preferential voting because they are less entrenched in the existing political system.

These findings suggest it is possible, with the continued implementation of preferential voting, that participation could increase and the political system could gain increased legitimacy amongst historically disenfranchised groups. A switch to preferential voting could increase the perceived political efficacy of minority voters, younger voters, and low-income voters; consequently, political participation amongst these groups could increase. As the analysis demonstrated, however, there are some groups that do not support preferential voting. The most surprising finding along these lines was the negative relationship between females and support for preferential voting. This relationship is the opposite to that hypothesized and, to the best of my knowledge, has no obvious theoretical explanation. I originally suspected the result was due to some mediator or interaction variable unaccounted for in the original model, but extensive exploration of those possibilities revealed the relationship to be consistent and robust. It is possible some mediator unaccounted for in the data set might yet explain that relationship, but as the data used here are the most comprehensive available it is unlikely further empirical light is going to be shed on this finding without additional data collections. After extensive probing of this relationship it remains something of a mystery and must be left to future research to pursue further.

It is difficult to say how the implementation of preferential voting could impact groups who do not support preferential voting in the long term; however, it should be

noted that none of the cities with preferential voting have experienced any sort of significant decline in political participation since abandoning plurality voting.

Chapter Three potentially represents this dissertation's most significant contribution. This chapter is significant not just because of the findings (which were mixed), but because of the method used to arrive at these findings. This chapter described a sophisticated program designed to collect and analyze tweets pertaining to a specific word, a collection of words, or hashtags. The program can perform various analyses on the tweets it collects, including a sentiment analysis, a readability analysis, and an image analysis for the purpose of facial recognition and classification. The inclusion of facial analysis represents a significant departure from previous Twitter scraper programs. While previously developed programs collect tweets, the analyses by this program allow for greater depth than any program previously.

The analysis of tweets related to Maine's preferential voting decision indicated that the majority of tweets were positive, which is consistent with the majority support the state's preferential voting ballot initiative received. The analysis demonstrated that, while some of the findings may not have reached significance, the socio-demographic characteristics which predicted support for preferential voting mirrored the conclusions found in Chapter Two in terms of directionality, but not significance. Specifically, younger Twitter users were significantly more likely to tweet positively regarding Maine's Question Five than their older counterparts. This indicates that similar to what I discovered in Chapter Two and Chapter Five, young people tend to support preferential voting as an alternative to traditional plurality voting. Additionally, this analysis found that there was an active and dedicated online effort committed to the implementation of

preferential voting at the state level. The Twitter scraper program can continue to be leveraged as a means of examining the future of preferential voting in the United States.

Chapter Three's method provides an extremely valuable tool for public opinion research. The use of Twitter allowed me to collect a tremendous number of people's opinions about a specific topic in a short amount of time for a small amount of money. This analysis has demonstrated that this program is useful for evaluating public opinion on an aggregate scale. The information gathered through this tool could have far-reaching implications for the study of public opinion. For example, I have independently verified that the information collected through this program can be used to accurately measure fluctuations in the presidential approval numbers collected by Gallup. Again, like the analysis in Chapter Three, Twitter was often biased because of its non-representative user base. As a result, day-to-day presidential approval numbers were typically over-inflated by a few points. Additionally, the scraper program used in this analysis can be used to collect information from respondents all over the world. While it may be difficult or expensive for a researcher to survey South Americans regarding their opinions of Donald Trump, the program used in Chapter Three is able to accomplish this feat for free and in a short amount of time. The sentiment analysis revealed that the majority of tweets pertaining to preferential voting, ranked choice voting, or Question Five were positive in nature, which reflects the electoral outcome in Maine. Furthermore, while the demographic information collected by the image analysis ultimately failed to reach statistical significance, the directionality of the correlation coefficients mirrored the findings of Chapter Two.

The program used in Chapter Three is useful, but as the discussion in that chapter

highlights, it can hardly be considered a magic bullet for public opinion research. Twitter is an extremely biased sample and can hardly be considered representative of the general electorate. The program described in Chapter Three would be most useful for an exploratory analysis or a supplementary analysis. Twitter would be most useful when a researcher would like to explore public opinion on a particular topic as a means of attaining funding for a traditional survey. Twitter would also be useful when a traditional survey cannot provide sufficient data, and a larger sample size would be the most useful.

This method should not be used for niche topics or topics that people may not have tweeted about. For example, it would be difficult to find people who may have posted about their concerns about a local bond issue or the impending extinction of the Iowa Pleistocene snail (Discus macclintocki). Still, despite its limitations, Twitter is a valuable tool for researchers who keep its limitations in mind. Additional information pertaining to the program (including Github repository URL) can be found in the appendix.

Chapter Four focused on the impact of preferential voting on the electoral process. Specifically, I examined whether the transition from plurality voting to preferential voting was accompanied by a reduction in perceptions of campaign negativity and a change in the campaign tactics used to reach voters. This chapter demonstrated that respondents in cities with preferential voting are likely to perceive lower levels of campaign negativity than respondents in cities with plurality voting. Additionally, respondents in cities with preferential voting were less likely to indicate hearing candidates criticize each other.

A reduction in negative campaigning could have significant positive implications

for the electoral process. As mentioned in Chapter Two, the scholarship regarding the electoral effectiveness of negative campaigning is divided. While the effectiveness of negative campaigning is debatable, it is difficult to dispute the fact that the majority of Americans are dissatisfied with the levels of negativity present in modern elections. The switch to preferential voting and the subsequent reduction in campaign negativity may not increase voter turnout, but it is likely to increase satisfaction with the political process. Additionally, the transition to preferential voting may increase the perceived legitimacy of elections and reduce polarization through the reduction in attack advertising.

These findings indicate that candidates in cities with preferential voting are less likely to rely on negative campaigning or attack advertising as a means of achieving electoral success. The second portion of this analysis demonstrated that the campaign tactics utilized in plurality systems are different than the tactics utilized in the preferential system. Candidates running for election in cities with preferential voting are more likely to utilize campaign tactics which engage and mobilize the voter personally. This finding demonstrates that there are significant differences between preferential systems and plurality systems regarding the effectiveness of various campaign tactics.

Chapter Five addressed one of the most common criticisms of preferential voting, that many voters lack the political knowledge or sophistication to engage with preferential voting. Chapter Five demonstrated that, despite its benefits, the transition to preferential voting may put some people at a disadvantage. The analysis found that education and gender were significant predictors of a respondent's ability to understand preferential voting. Respondents who had less education were female, and were

unemployed were less likely to understand preferential voting. Furthermore, minority respondents and female respondents were less likely to understand the instructions they were required to read prior to engaging in the preferential voting process. However, it should be noted that the socio-demographic characteristics which predict difficulty with preferential voting were not too different than the socio-demographic characteristics which predict difficulty with plurality voting. That said, there is no doubt the empirical findings in Chapter Five do support at least some of the points made by critics of preferential voting. Ultimately, preferential voting may need to achieve greater mainstream recognition until it can be better understood by historically disenfranchised groups.

The majority of respondents in cities with preferential voting have a strong desire to see preferential voting persist in their cities, so it is unlikely that these cities will revert to plurality voting. Additionally, Maine's recent successful adoption of preferential voting will provide a positive example for other states, encouraging them to adopt preferential voting as a means of conducting state and local elections.

One of the most pervasive findings in this project, however, was that female respondents consistently rejected preferential voting. Specifically, as Chapter Two demonstrated, women in cities with preferential voting were statistically likely to reject its continued implementation. These findings were further backed up by Chapter Three, which demonstrated that women were less likely to tweet positively about preferential voting. Unfortunately, this finding does not appear to be tied into any specific variables which appeared in the survey used for this project. I suspect that these results may be the result of overpromises made during the campaigns to ratify preferential voting, due to

sampling issues or are due to an interaction for which I cannot account.

This project has touched on several important questions pertaining to preferential voting in the United States. However, the answers produced by research should ultimately be used to make a prediction regarding the future of preferential voting in the United States. Preferential voting will likely continue its growth in the United States at the local and state level, but for the time being, preferential voting is being adopted primarily in left-leaning parts of the country. Maine became the first state to adopt preferential voting at the state level, but its implementation was fraught with difficulty and political obstruction. Preferential voting represents a shift away from the two-party system, which has allowed many politicians to amass tremendous social and political capital. Maine's Governor LePage, for example, has benefitted tremendously from Maine's previous system of plurality voting, which might explain his animosity towards preferential voting. Because of these challenges, it is unlikely that preferential voting will be adopted by states with deeply entrenched political factions. The future of preferential voting at the state level is likely to be influenced by Maine's ability to implement preferential voting and the electorate's reception to their newly implemented voting system.

Preferential voting would alleviate many of the electoral woes currently plaguing our national system of politics, but the implementation of preferential voting at the national level would require a constitutional amendment. If the fight over Maine's recent referendum on preferential voting is an indication, implementing preferential voting at the federal level would be impossibly challenging. Preferential voting would seriously undermine the power of the two-party system currently dominating American politics.

Through the increased choice it provides voters, preferential voting would allow for political options outside of the Democrat and Republican dichotomy to which many voters have become accustomed. While plurality all but guarantees a two-party system, preferential voting allows third-party candidates to have a better chance of winning. It is unlikely that, despite all the benefits associated with preferential voting, either political party would advocate for the implementation of a voting system that would undermine their political power.

Certainly, a voting system that works for one city may be woefully ill-equipped to handle the demands of another. This research has demonstrated that voting systems vary in their utility in specific use cases. Ultimately, some cities and states may want to continue with plurality voting because it is easy to understand and is well-suited for a less diverse political environment. Conversely, some cities may choose to adopt preferential voting because it has less negativity, is popular among specific socio-demographic groups, and caters to a diverse political environment. It is beholden on a city, state, or country to examine the benefits of each to decide if preferential voting is the right decision.

Works Cited

Abramowitz, A. I., & Webster, S. (2016). The rise of negative partisanship and the nationalization of US elections in the 21st century. *Electoral Studies*, *41*, 12-22.

Ansolabehere, S., & Iyengar, S. (1996). Can the press monitor campaign advertising? An experimental study. *Harvard International Journal of Press/Politics*, *1*(1), 72-86.

Arrow, K. J., & Raynaud, H. (1986). Social choice and multicriterion decision- making. *MIT Press Books*, *1*.

Bartels, L. M. (2016). *Unequal democracy: The political economy of the new gilded age*. Princeton University Press.

Babcock, L., & Laschever, S. (2009). *Women don't ask: Negotiation and the gender divide*. Princeton University Press.

Bartholdi, J. J., & Orlin, J. B. (1991). Single transferable vote resists strategic voting. *Social Choice and Welfare*, *8*(4), 341-354.

Beauchamp, N. (2017). Predicting and interpolating state-level polls using Twitter textual data. *American Journal of Political Science*, *61*(2), 490-503.

Bhat, C. R. (1994). Imputing a continuous income variable from grouped and missing income observations. *Economics Letters*, *46*(4), 311-319.

Blais, A., Aldrich, J. H., Indridason, I. H., & Levine, R. (2006). Do voters vote for government coalitions? Testing Downs' pessimistic conclusion. *Party Politics*, *12*(6), 691-705.

Blake, A. (2016, July 29). Welcome to the next, most negative presidential election of our lives. Retrieved from https://www.washingtonpost.com/news/the-fix/wp/2016/07/29/clinton-and-trump-accept-their-nominations-by-telling-you-what-you-should-vote-against/?noredirect=on&utm_term=.d2cd7dda3b0e

Bowler, S., Donovan, T., & Brockington, D. (2003). *Electoral reform and minority representation: Local experiments with alternative elections*. Ohio State University Press.

Brader, T. (2005). Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions. *American Journal of Political Science*, *49*(2), 388-405.

Chandra, K. (2005). Ethnic parties and democratic stability. *Perspectives on Politics*, *3*(2), 235-252.

Collins, P. H. (2002). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. Routledge.

Cook, C., & Latterman, D. (2011). Ranked Choice Voting in the 2011 San Francisco Municipal Election.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. John Wiley & Sons.

Donovan, T., Tolbert, C., & Gracey, K. (2016). Campaign civility under preferential and plurality voting. *Electoral Studies*, *42*, 157-163.

Earl, J., & Kimport, K. (2011). *Digitally enabled social change: Activism in the internet age*. Mit Press.

Farquharson, R. (1969). *Theory of voting*. Blackwell.

Fowler Jr, F. J. (2013). *Survey research methods*. Sage publications.

Fraenkel, J., & Grofman, B. (2006). Does the alternative vote foster moderation in ethnically divided societies? The case of Fiji. *Comparative Political Studies*, *39*(5), 623-651.

Fridkin, K. L., & Kenney, P. J. (2004). Do negative messages work? The impact of negativity on citizens' evaluations of candidates. *American Politics Research*, *32*(5), 570-605.

Geiger, A. (2016, September 02). For many voters, it's not which presidential candidate they're for but which they're against. Retrieved from http://www.pewresearch.org/fact-tank/2016/09/02/for-many-voters-its-not-which- presidential-candidate-theyre-for-but-which-theyre-against/

Gerber, A. S., Green, D. P., & Shachar, R. (2003). Voting may be habit-forming: evidence from a randomized field experiment. *American Journal of Political Science*, *47*(3), 540-550.

Green, D. P., & Gerber, A. S. (2015). *Get out the vote: How to increase voter turnout*. Brookings Institution Press.

Haan, N., Smith, M. B., & Block, J. (1968). Moral reasoning of young adults: Political-social behavior, family background, and personality correlates. *Journal of Personality and Social Psychology*, *10*(3), 183.

Herrnson, P. S., Lay, J. C., & Stokes, A. K. (2003). Women running "as women": Candidate gender, campaign issues, and voter-targeting strategies. *The Journal of Politics*, *65*(1), 244-255.

Herrnson, Paul S., and Jennifer C. Lucas. "The fairer sex? Gender and negative campaigning in US elections." *American Politics Research* 34, no. 1 (2006): 69-94.

Hetherington, M. J., & Hetherington, M. J. (2005). *Why trust matters: Declining political trust and the demise of American liberalism*. Princeton University Press.

Hetherington, M. J., & Rudolph, T. J. (2015). *Why Washington won't work: Polarization, political trust, and the governing crisis*. University of Chicago Press.

Hibbing, J. R., & Theiss-Morse, E. (2002). *Stealth democracy: Americans' beliefs about how government should work*. Cambridge University Press.

Horowitz, M. J. (1985). Disasters and psychological responses to stress. *Psychiatric Annals*, *15*(3), 161-167.

Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. *Public opinion quarterly*, *68*(1), 94-101.

Krupnikov, Y. (2011). When does negativity demobilize? Tracing the conditional effect of negative campaigning on voter turnout. *American Journal of Political Science*, *55*(4), 797-813.

Lau, R. R., & Pomper, G. M. (2004). *Negative campaigning: An analysis of US Senate elections*. Rowman & Littlefield.

Lau, R. R., Sigelman, L., & Rovner, I. B. (2007). The effects of negative political campaigns: a meta-analytic reassessment. *Journal of Politics*, *69*(4), 1176-1209.

Lee, D. S., Moretti, E., & Butler, M. J. (2004). Do voters affect or elect policies? Evidence from the US House. *The Quarterly Journal of Economics*, *119*(3), 807-859.

Link, M. W., Battaglia, M. P., Frankel, M. R., Osborn, L., & Mokdad, A. H. (2007). Reaching the US cell phone generation: comparison of cell phone survey results with an ongoing landline telephone survey. *Public Opinion Quarterly*, *71*(5), 814-839.

Lui, C., Metaxas, P. T., & Mustafaraj, E. (2011). On the predictability of the US elections through search volume activity.

Maccoby, E. E. (1954). Why do children watch television? *Public opinion quarterly*, *18*(3), 239-244.

McElroy, G., & Marsh, M. (2010). Candidate gender and voter choice: Analysis from a multimember preferential voting system. *Political Research Quarterly*, *63*(4), 822-833.

Mellon, J., & Prosser, C. (2017). Missing Nonvoters and Misweighted Samples: Explaining the 2015 Great British Polling Miss. *Public Opinion Quarterly*, *81*(3), 661-687.

Michelson, M. R. (2003). Getting out the Latino vote: How door-to-door canvassing influences voter turnout in rural central California. *Political Behavior*, *25*(3), 247-263.

Mistler, S. (2018, June 10). Maine Voters To Use Ranked-Choice Voting. Retrieved from https://www.npr.org/2018/06/10/618653587/maine-voters-to-use-ranked-choice-voting

Neely, F., Blash, L., & Cook, C. (2005). *An assessment of ranked-choice voting in the San Francisco 2004 election*. Public Research Institute, San Francisco State University.

Nilsen, E. (2018, June 12). Maine's governor is threatening not to certify the results of Tuesday's primary. Retrieved from https://www.vox.com/2018/6/12/17453676/paul-lepage-maine-governor-threatening- not-to-certify-tuesday-primary-results-2018-midterms

Norris, P. (2004). Young people & political activism. *Harvard University, John F. Kennedy School of Government. (32p).*

Obata, T., & Ishii, H. (2003). A method for discriminating efficient candidates with ranked voting data. *European Journal of Operational Research*, *151*(1), 233-237.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1-135.

Polsby, N. W. (1963). *Community power and political theory*(Vol. 7). Yale University Press.

Poundstone, W. (2008). *Gaming the vote: Why elections aren't fair (and what we can do about it)*. Macmillan.

Project FairVote. Retrieved from http://www.FairVote.org

Rabushka, A., & Shepsle, K. A. (1972). *Politics in plural societies* (pp. 208-212). Redwood City: Stanford University Press.

Reilly, B. (2001). *Democracy in divided societies: Electoral engineering for conflict management*. Cambridge University Press.

Rohe, W. M., & Stegman, M. A. (1994). The effects of homeownership: On the self-esteem, perceived control and life satisfaction of low-income people. *Journal of the American Planning Association*, *60*(2), 173-184.

Seelye, K. Q. (2017, May 23). Ranked-Choice Voting System Violates Maine's Constitution, Court Says. Retrieved from https://www.nytimes.com/2017/05/23/us/maine-ranked-choice-elections-voting.html

Shah, D. V., Cho, J., Nah, S., Gotlieb, M. R., Hwang, H., Lee, N. J., ... & McLeod, D. M. (2007). Campaign ads, online messaging, and participation: Extending the communication mediation model. *Journal of communication*, *57*(4), 676-703.

Tolbert, C., & John, S. (2015). Socioeconomic and Demographic Perspectives on Ranked Choice Voting in the Bay Area. FairVote: The Center for Voting and Democracy. Retrieved October 18, 2018

Traugott, M. W., & Kang, M. E. (2000). Push polls as negative persuasive strategies. *Election polls, the news media, and democracy*.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. *Icwsm*, *10*(1), 178-185.

Turrell, G. (2000). Income non-reporting: implications for health inequalities research. *Journal of Epidemiology & Community Health*, *54*(3), 207-214.

United States Census Bureau. 2010 Census.U.S. Census Bureau. 2010. Web. 18 October 2018

Weil, G. L. (2016, September 30). Ranked-choice voting: Costly, complicated, undemocratic. Retrieved from https://www.pressherald.com/2016/09/30/maine-voices-ranked-choice-voting-costly-complicated-undemocratic/

Wolfinger, R. E., & Rosenstone, S. J. (1980). *Who votes?* (Vol. 22). Yale University Press.

Appendix A: Chapter Three Program Details

Setup

Programs Needed: Aside from the necessary Python modules, which are described below, the program built for this analysis requires two programs: MongoDB and Python. Prior to running the program, the user is required to specify a database. This is accomplished by Installing mongoDB and running mongod.exe to specify a database path. The program is designed to work on Python 3.5 or 3.6. The packages utilized by the program may not work properly on other versions of Python. It is important to remember not to run mongod.exe on Windows through Ubuntu/SUSE/Fedora, as it will not work properly

Modules Needed: The program uses a variety of python modules. These modules can work on either Linux or Windows systems, but installation may vary between operating systems. *Tweepy* was used for tweet collection, *nltk* was used for text analysis, *requests* were used for image analysis, and *texts that* was used for readability analysis. The program used *pymongo* to communicate with the non-relational database and *python-levenshtein* was used for duplicate checking. Note for Windows users: *python-levenshtein* may fail to install, so you'll need to download a .whl file.

Usage: Twitter API keys can easily be acquired by setting up a developer account at Twitter's developer portal. Similarly, Kairos keys can be acquired by signing up for a Kairos account and creating an app. We recommend using MongoDB Compass or Robo 3T if you want a good visual view of your databases.

Data Collection: There are two approaches to data collection supported by the program: real-time and historical. This project used the program's abilities to collect and analyze historical tweets to draw broader conclusions pertaining to public support for preferential voting and to further test the conclusions reached in Chapter Two.

***Real-Time Collection*:** In this approach, the program searches for tweets based on search term(s) input by the user. The user also specifies the number of tweets collected. Tweets and their associated information are retrieved using the Twitter streaming API Python module, Tweepy. The tweets pulled from a single search are stored within a collection and are assigned a default name, which is the first search term and the date/time the search was initiated. These collections are stored in a non-relational database. The user can search tweets based on keywords or hashtags contained in tweets, or by individual Twitter users. Searching by an individual user does not search that user's entire Twitter history; instead, searching for an individual user will stream that user's most recent tweets in real-time.

One of the largest issues to overcome regarding the collection of tweets was the issue of duplicates. Duplicates occur when a single tweet is shared by multiple users. Duplicate tweets have become a large problem because of the prevalence of Twitter bots, programs designed to send out automated posts on Twitter. These bots often send out the same message from multiple accounts for the sake of advertising or to create a false appearance of ideological consensus. To address this, the program compares each tweet to other tweets in the collection for similarity. First, the program removes the punctuation and spaces from the current tweet as a means of making the comparison between tweets less resource-intensive. After this step, the program first checks if there are any identical

tweets in the database. The current tweet is then compared to every other tweet in the database using Python-Levenshtein's distance calculation to determine a similarity ratio between the two tweets, resulting in a decimal value. This similarity ratio is compared to the similarity threshold established by the user to determine if the tweet should be thrown out. If a tweet is found to be too similar to another tweet in the database, it is deemed a duplicate. If a duplicate is found, the two tweets are compared, and the tweet with the most favorites is kept. If the two tweets have the same number of favorites, the older tweet is kept. This ensures that the program collects the maximum amount of data without compromising the integrity of the data it collects. This duplicate-checking process ensures that the final dataset does not contain identical tweets. The duplicate checking process, however, does not filter out multiple unique tweets by a single user.

Multiple tweets from a single user could prove to be an issue due to some user's penchant for carpet-bombing Twitter with the same opinion expressed multiple ways. Multiple tweets from a single user are easy to remove during the data cleaning process, which is what I did for this analysis.

***Historical Collection*:** The program can also collect tweets from specific dates and times, and this process does not differ significantly from that of collecting real-time tweets described above. Historical tweet-gathering differs from streaming in that the user can specify a "before" date and an "after" date. These dates will ensure that the program collects tweets between the dates chosen by the user. Additionally, the user can set a "result type" which specifies the types of tweets gathered: "popular" only returns the most popular tweets, which is determined by the number of favorites a given tweet

receives, “recent” only returns the most recent tweets, and “mixed” (the default) returns a combination of recent and popular tweets to ensure a representative sample.

Tweets are gathered 100 at a time until Twitter requires fewer tweets to be gathered based on the API limit, which is established by Twitter. In this context, an API limit is the amount of data Twitter will allow you to collect within a specific timeframe. When the user has exceeded the API limit, the program will then pause its collection until Twitter allows it to resume. These tweets are then filtered using the above methods. Successful tweets are counted and inserted into the database. It should be noted that throttling can occur in the real-time collection; however, it is more common for the collection to throttle during historical collection given that tweets can be collected significantly faster during historical collection.

Sentiment Analysis:

Regardless of whether the data collection is historical or conducted in real time, once the tweets have been collected they are subjected to a sentiment analysis. The sentiment analysis relies on the Natural Language Toolkit (NLTK) module in Python. When creating the program, we trained NLTK to understand tweet polarity using the Vader Sentiment lexicon (Hutto & Gilbert 2014). This lexicon is specifically designed to evaluate the emotional polarity of social media posts. Like most types of sentiment analysis, ours utilizes a naive Bayes classifier in conjunction with a series of training sets derived from the Vader lexicon. A sentiment analysis can be run on either a collection or a database of tweets. The program scores each tweet based on its emotional polarity; tweets are either classified as positive, neutral, or negative. These classifications are added to each tweet’s corresponding database entry. Each tweet is also given a compound

score, which is computed by combining the scores of each word in the lexicon and then normalizing it to a value between -1 and +1. The compound score represents a single unidimensional measure for the emotional polarity of a given tweet.

Facial Analysis:

The next step performed by the program is facial recognition and classification. This step is optional; if a user only wants to determine whether or not the tweet content is positive or negative, the sentiment analysis is sufficient. The facial recognition and classification process are designed to identify each Twitter user's specific demographic characteristics and tie those characteristics to that user's tweet. The Twitter user's display picture is collected and submitted to a facial recognition application called Kairos. Kairos has been integrated into the program's analysis protocol to provide the best possible method of classifying users based on their age, ethnicity, and gender. Other companies, such as Microsoft, Amazon, and IBM, have created facial recognition applications; however, none of these programs is able to classify faces based on ethnicity.

Kairos can also identify the emotion being displayed on the face of the individual in the profile picture and whether the individual is wearing glasses. It should be noted that, if the profile does not contain a profile picture with a single identifiable human face, the program disregards that particular image and continues with the analysis. Similarly, if the analysis is unable to identify the user's age, gender, and ethnicity with a high degree of confidence, that image is excluded from the analysis. Finally, Kairos also excludes any users' display pictures in GIF format, since Kairos does not support this file format. In terms of accuracy, Kairos claims to provide accuracy exceeding its competitors, such as Amazon, Microsoft, and IBM. Ultimately, it is difficult to measure this claim, as Kairos

keeps its more specific methods a secret; however, it is important to note that accuracy in facial recognition is influenced by a variety of variables. Thankfully, Kairos includes confidence ratings which allow for the researcher to quickly eliminate data points which may negative impact a study's internal validity.

Readability:

The program can also assess the reading complexity of the collected tweets. The readability analysis uses the python module "textstat" to simplify the process of finding the various readability values. This process analyzes the individual tweets to determine three separate readability values: Flesch Reading Ease, Flesch–Kincaid Grade Level, and standardized reading score. The appropriate scores are inserted into the database with the tweet. Measuring the readability of a tweet is difficult, however, and sometimes the readability scores produced by the program cannot be used for research purposes because the tweet is too short or does not contain enough usable words. If a tweet cannot be reliably used for research purposes, it is not inserted into the database. This feature was originally added to the program to serve as a proxy measurement for education. Unfortunately, the readability feature was also created and implemented with Twitter's new 280- character limit in mind. While Flesch-Kincaid scores are a reasonable proxy measurement for someone's level of education, the 140-character limit in place at the time these tweets were collected adds too great off a confound to use these readability scores in this analysis.

CSV Export:

As described above, all of the collected and analyzed data are stored in a database. This format does not lend itself to easy analysis, so the program is equipped with an

export feature which can either export a single collection or an entire database, into a convenient comma-separated values file (CSV). Comma-separated values files are compatible with most statistical packages and allow for convenient analysis of the collected data.

Ultimately, the result of the web scraper program just described is a spreadsheet with each line containing a unique tweet, the emotional polarity of that tweet, and the demographic indicators described above. This spreadsheet is incredibly useful for this analysis in that it allows for easy analysis of tweets. This information allows for the examination of the ratio of positive to negative tweets about a given topic as well as information about what kinds of people are tweeting positively or negatively about that topic. This approach is similar to traditional survey research in that it allows the researcher to gather information about people's opinions and demographic information; however, unlike survey research, this approach is free, incredibly fast, and can collect data from millions of respondents.

Appendix B: Tables

Table 2.1: Descriptive Statistics	Col %	No.
Local Elections should use Preferential Voting		
Yes	41.86	1028
No	58.14	1428
Total	100.0	2456
City Type		
Uses Preferential Voting	54.76	1345
Does Not Use Preferential Voting	45.24	1111
Total	100.0	2456
Gender		
Male	44.42	1091
Female	55.58	1365
Total	100.0	2456
Race		
White	55.21	1356
Nonwhite	44.79	1100
Total	100.0	2456
Partisanship		
Republican	12.50	298
Democrat	64.05	1527
Independent	23.45	559
Total	100.0	2384
Income		
Low Income	34.43	744
High Income	41.37	894
Missing Income	24.20	523
Total	100.0	2161

Source: 2014 Project FairVote

Table 2.2 Predicting Support for Preferential Voting

	Dependent variable:			
	Support Use of Preferential Voting			
	(1)	(2)	(3)	(4)
Preferential City	0.347***	0.311***	0.332***	0.294***
	(0.088)	(0.089)	(0.088)	(0.089)
Age	−0.018***	−0.018***	−0.025***	−0.026***
	(0.002)	(0.002)	(0.004)	(0.004)
White	−0.135	−0.163*	−0.146	−0.172*
	(0.090)	(0.091)	(0.090)	(0.091)
Democrat	−0.160		−0.134	
	(0.100)		(0.099)	
Republican	−0.363**		−0.354**	
	(0.153)		(0.152)	
Party(Democrat-Republican)		−0.033		−0.045
		(0.063)		(0.063)
Low Income	0.295**	0.299**		
	(0.146)	(0.147)		
High Income	−0.024	−0.042		
	(0.144)	(0.144)		
Missing INcome	−0.265*	−0.285*		
	(0.156)	(0.158)		
Imputed Income			−0.216**	−0.244**
			(0.102)	(0.104)
Education	−0.095***	−0.083**	−0.019	0.005
	(0.032)	(0.033)	(0.058)	(0.060)
Female	−0.232***	−0.252***	−0.310***	−
	0.340*** (0.087)		(0.088)	(0.097)
	(0.098)			
Married	0.276***	0.249***	0.516***	0.525***
	(0.091)	(0.093)	(0.172)	(0.175)
Employed	−0.004	−0.003	0.080	0.090
	(0.085)	(0.086)	(0.090)	(0.091)
Constant	1.240***	1.141***	2.059***	2.070***
	(0.269)	(0.285)	(0.372)	(0.388)
Observations	2,436	2,369	2,437	2,370
Log Likelihood	−1,587.364	−1,548.276	−1,595.642	−1,556.548
Akaike Inf. Crit.	3,200.727	3,120.551	3,213.285	3,133.095

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.1: Descriptive Statistics	Col %	US Population %	No.	% Pos.
Gender				
Female	41.60	50.9	4988	63.52
Male	58.40	49.1	7001	63.92
Total	100.0	100	11989	
Age Group				
Young Adult	73.87	32	5864	64.52
Adult	14.50	39.45	1151	62.41
Senior	11.63	16.63	923	68.68
Total	100.0		8273	
Ethnicity				
Asian	8.77	4.8	929	61.76
Black	10.24	12.6	1085	62.08
Hispanic	11.02	16.3	1167	67.83
Other	4.48	6.2	475	58.24
White	65.49	72.4	6937	63.98
Total	100.0		10593	

Source: 2014 Project FairVote

Table 3.2

	Dependent variable:
	Tweeted Positively
Age	-0.007**
	(0.044)
White	0.055
	(0.429)
Male	0.764
	(0.759)
Observations	2508
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4.1: Perceptions of Negative Campaigning

	Dependent variable:			
	Negativity of Election Compared to Previous Elections			
	(1)	(2)	(3)	(4)
Preferential City	1.393***	1.384***	1.400***	1.391***
	(0.094)	(0.096)	(0.094)	(0.096)
Age	−0.010***	−0.010***	−0.009**	−0.009**
	(0.002)	(0.002)	(0.004)	(0.004)
White	0.078	0.054	0.082	0.059
	(0.090)	(0.091)	(0.089)	(0.091)
Democrat	0.020		0.046	
	(0.099)		(0.099)	
Republican	0.188		0.205	
	(0.146)		(0.146)	
Party (Democrat-Republican)		0.062		0.053
		(0.061)		(0.061)
Low Income	0.026	0.031		
	(0.144)	(0.146)		
High Income	0.169	0.188		
	(0.143)	(0.144)		
Missing Income	−0.098	−0.094		
	(0.151)	(0.154)		
Imputed Income			0.063	0.063
			(0.101)	(0.103)
Education	−0.017	−0.017	−0.036	−0.035
	(0.032)	(0.032)	(0.058)	(0.059)
Female	0.056	0.060	0.067	0.070
	(0.086)	(0.087)	(0.096)	(0.097)
Married	0.072	0.075	0.017	0.025
	(0.089)	(0.090)	(0.170)	(0.173)
Employed	−0.063	−0.050	−0.080	−0.067
	(0.084)	(0.085)	(0.089)	(0.090)
Observations	2,281	2,225	2,282	2,226

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.2: Candidate Campaign Behavior (Praise)

	Dependent variable:			
	Witnessed Candidates Praising Each Other			
	(1)	(2)	(3)	(4)
Preferential City	0.098	0.088	0.097	0.087
	(0.089)	(0.090)	(0.089)	(0.091)
Age	0.001	0.001	−0.003	−0.003
	(0.002)	(0.002)	(0.004)	(0.004)
White	−0.310***	−0.299***	−0.308***	−
	0.297***	(0.092)	(0.093)	(0.092)
	(0.093)			
Democrat	−0.013		−0.007	
	(0.102)		(0.102)	
Republican	−0.020		−0.007	
	(0.153)		(0.153)	
Party (Democrat-Republican)		−0.008		−0.006
		(0.064)		(0.063)
Low Income	0.189	0.218		
	(0.150)	(0.152)		
High Income	0.146	0.151		
	(0.149)	(0.150)		
Missing Income	0.072	0.097		
	(0.159)	(0.161)		
Imputed Income			−0.135	−0.116
			(0.103)	(0.106)
Education	0.032	0.020	0.092	0.069
	(0.034)	(0.034)	(0.060)	(0.061)
Female	0.087	0.051	0.027	0.002
	(0.089)	(0.090)	(0.099)	(0.100)
Married	−0.003	0.001	0.189	0.161
	(0.092)	(0.094)	(0.175)	(0.179)
Employed	0.143*	0.155*	0.187**	0.192**
	(0.086)	(0.088)	(0.092)	(0.094)
Observations	2,302	2,242	2,303	2,243

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.3: Candidate Campaign Behavior (Criticize)

	Dependent variable:			
	Witnessed Candidates Criticizing Each Other			
	(1)	(2)	(3)	(4)
Preferential City	−1.219***	−1.221***	−1.205***	−
	1.210***	(0.085)	(0.086)	(0.085)
	(0.086)			
Age	0.002	0.002	0.007*	0.006
	(0.002)	(0.002)	(0.004)	(0.004)
White	−0.155*	−0.162*	−0.144*	−0.151*
	(0.085)	(0.086)	(0.085)	(0.085)
Democrat	−0.008		−0.011	
	(0.095)		(0.094)	
Republican	−0.270*		−0.262*	
	(0.142)		(0.141)	
Party (Democrat-Republican)		−0.088		−0.083
		(0.059)		(0.059)
Low Income	−0.120	−0.125		
	(0.135)	(0.136)		
High Income	0.158	0.131		
	(0.132)	(0.133)		
Missing Income	0.155	0.158		
	(0.142)	(0.144)		
Imputed Income			0.175*	0.151
			(0.096)	(0.097)
Education	0.058*	0.059*	−0.0005	0.010
	(0.031)	(0.031)	(0.055)	(0.055)
Female	−0.157*	−0.173**	−0.096	−0.121
	(0.081)	(0.082)	(0.091)	(0.092)
Married	0.009	−0.004	−0.182	−0.165
	(0.085)	(0.086)	(0.162)	(0.165)
Employed	0.111	0.086	0.053	0.035
	(0.079)	(0.080)	(0.084)	(0.085)
Observations	2,150	2,103	2,151	2,104

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.4 Anova Results: Comparing Campaign Activities (In-Person Contact)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Between Groups	1	1.21874	1.21874	4.88	0.0272
Within Groups	1895	472.920	.24956		
Total	1896	474.139	.2500		

Preferential City	Mean	Std. Dev.	Frequency
No	.4644	.49902	859
Yes	.51541	.50000	1038
Total	.49235	.50007	1897

Table 4.5 Anova Results: Comparing Campaign Activities (Cumulative Campaign Contact)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Between Groups	1	10.589	10.589	9.48	0.0021
Within Groups	1818	1.11759	.24956223		
Total	1819	2042.37	1.12279		

Preferential City	Mean	Std. Dev.	Frequency
No	2.6450	1.0673	834
Yes	2.7981	1.0484	986
Total	2.728	1.059	1820

Table 5.1

	Dependent variable:		
	Understand Preferential	Understand Plurality	Understand Instructions
	(1)	(2)	(3)
Democrat	−0.006	0.130	0.035
	(0.120)	(0.140)	(0.107)
Republican	−0.237	−0.036	0.197
	(0.214)	(0.183)	(0.161)
Low Income	−0.001	−0.219	−0.220
	(0.176)	(0.191)	(0.157)
High Income	0.401**	−0.269	0.016
	(0.171)	(0.194)	(0.157)
Missing Income	−0.048	−0.500**	−0.179
	(0.185)	(0.203)	(0.165)
Education	0.101**	0.200***	0.048
	(0.040)	(0.044)	(0.035)
White	0.168	0.092	0.233**
	(0.107)	(0.129)	(0.096)
Female	−0.342***	−0.115	−0.205**
	(0.104)	(0.120)	(0.093)
Age	0.0002	0.009***	−0.001
	(0.003)	(0.003)	(0.003)
Married	0.136	0.230*	−0.134
	(0.108)	(0.125)	(0.097)
Employed	−0.182*	−0.272**	−0.142
	(0.101)	(0.121)	(0.091)
Observations	1,322	988	2,159

Note: *p<0.1; **p<0.05; ***p<0.01