Mudslinging at the Margins: The Conditionality of Negative Campaigning Under Ranked-Choice Voting in the United States

DARTMOUTH

By
Mitchell Ransden

Honors Thesis
Professor Jennifer Jerit, Advisor
Department of Government
Dartmouth College
April 28, 2023

For Jonathan, Alex, Steve, and the rest of the team at Voter Choice CT

Contents

Acknowledgements iv

Abstract viii

Introduction: A more moderate system? 1

Chapter 1: A new ranked-choice voting dataset 25

Chapter 2: Testing the conditionality of the "moderating incentive" of RCV 45

Chapter 3: The Maine effects: case studies of campaign rhetoric 77

Concluding Material 101

Bibliography 105

Acknowledgements

This thesis began as a harebrained project to answer my unresolved questions about electoral systems and campaign strategy. While the final text is deeply personal, it never would have come to fruition without the help of many talented individuals, all of whom have my deepest thanks.

Yuki Atsusaka has believed in the success of this project even more than I have, and his knowledge of the current landscape of ranked-choice voting research was invaluable to selecting and refining my research question. Wendel Cox provided further assistance with respect to digging through contemporary research, and together we made sure that these questions were, indeed, unanswered. My research might have remained a term paper were it not for the encouragement and support of Brendan Nyhan and Kathleen Powers, who gave me the chance to learn the ropes of political science research and persuaded me to take a big academic swing.

Simon Stone made time to meet with me as I learned the ropes of speech-related data collection and text analysis. Without his insights, I would still be lost in a forest of municipal debates and decade-old Twitter posts. I thank Breeze Floyd and the rest of the staff at the Wesleyan Media Project for graciously providing supplemental television advertising data. A generous grant from the Rockefeller Center for Public Policy and the Social Sciences allowed me to collect and analyze data using the latest and greatest in audio processing tools, and I thank them for their support.

The empirical portions of this thesis would not have been possible without the extensive help of Charles Crabtree. He provided insight into modeling strategies, sentiment analysis methods, and proper statistical methods on a continuing basis. Further, he took time out of his very busy schedule to read over initial drafts of different chapters well before becoming my

second reader. Any novel statistical insights are his, and the errors are mine.

The honors thesis program offered by the Department of Government is rigorous, and its high standards are maintained through the tireless effort of dedicated individuals. Roberta Seaver undertook the daunting task of providing background and scheduling support for all ten theses, and I thank her for creating an opportunity to present my work. Lucas Swaine served as co-director of the Honors Program, and his experience and good humor kept us all on the straight and narrow path of success. Other than my advisor, Benjamin Valentino spent more time than anyone working with me. Over the course of many months, face-to-face meetings, and sets of comments, he helped me refine my theoretical approach, overcome obstacles, and effectively communicate my results. Lastly, I thank the other members of "subgroup Valentino" for their thoughtful comments throughout this process and congratulate them on their success.

Jennifer Jerit was the best advisor any student could ask for. She kept me on track through every step of the process, read over every draft, and provided swift feedback even when at her busiest. It is because of her patience and painstaking line edits that this thesis is readable at all. She went above and beyond in every way and always pushed me to surpass my own expectations. For that and so much more, I cannot thank her enough.

Last but certainly not least, I thank my friends (some of whom had theses of their own!) and my family. They provided support and a healthy amount of distraction when I needed it most. Best of all, they listened to my endless ramblings on the subject of ranked-choice voting with boundless patience. Being able to rely on them throughout the last several months has been both relief and a delight and made my research better as a result.

List of tables and figures

Table 1.1: Statement search and text processing protocols 30

Figure 1.1: Distribution of RCV elections by year 34

Figure 1.2: Number of candidates in sampled elections 35

Figure 1.3: Distribution of ranking restrictions in RCV elections 36

Figure 1.4: Historical competitiveness of RCV elections 37

Table 1.2: Distribution of additional candidate-level and election-level attributes 38

Figure 1.5: Distribution of tweets among RCV candidates 39

Figure 1.6: Proportion of RCV candidates that use Twitter per year 39

Figure 1.7: Distribution of debate statements in RCV elections 41

Table 1.3: RCV campaign messages by type, year, and election 42

Figure 2.1: Multilevel model structure 49

Table 2.1: Examples of campaign statement sentiment scoring 55

Table 2.2: Correlation coefficients of LIWC and VADER sentiment measures 58

Table 2.3: LIWC sentiment ratings by text type 59

Figure 2.2: Candidate-level variables and campaign sentiment 60

Figure 2.3: Incumbent presence and campaign sentiment 61

Figure 2.4: Ranking restrictions and campaign sentiment 62

Figure 2.5: Number of candidates and campaign sentiment 63

Figure 2.6: Logged number of candidates and campaign sentiment 63

Figure 2.7: Historical competitiveness and campaign sentiment 64

Figure 2.8: Gender and incumbency interaction 65

Table 2.4: Multilevel models of campaign sentiment 66

Figure 2.9: Effects of candidate and electoral features on campaign sentiment (Model 1)..... 68

Figure 2.10: Effects of candidate and electoral features on campaign sentiment (Model 2)..... 68

Figure 2.11: Effects of candidate and electoral features on campaign sentiment (Model 3)..... 69

Figure 3.1: News articles that mention RCV by year 78

Table 3.1: Characteristics of federal RCV elections..... 79

Figure 3.2: Types of ads in the 2020 ME Senate race 91

Figure 3.3: Map of Maine congressional districts 93

Figure 3.4: Types of ads in the 2018 ME Second Congressional District race 96

Abstract

An increasing number of U.S. states and municipalities have replaced plurality voting with ranked-choice voting. Politicians and RCV advocates often claim that candidates in RCV elections are incentivized to use less negative campaigning. However, studies of negativity in RCV elections have only analyzed a handful of cases and fail to show the generalizability this "moderating incentive." To resolve these shortcomings, I compile a new comprehensive dataset of all major RCV elections in the United States and analyze more than 38,000 Twitter posts and statements made during candidate debates. Using sentiment analysis and three qualitative case studies, I show that rhetorical moderation under RCV is conditioned by candidate-level factors (gender and incumbent status) as well as election-level factors (incumbent presence). Furthermore, I provide preliminary evidence that the moderating incentive of RCV may be robust across electoral contexts.

Introduction: A more moderate system?

The United States is undergoing a major shift in electoral administration. Since 2004, a growing contingent of cities, counties, and states have replaced a plurality, "first-past-the-post" system of picking candidates with ranked-choice voting (RCV). Proponents of RCV claim that it is a remedy for numerous problems endemic to modern campaigns (FairVote 2023; League of Women Voters 2020), including elevated levels of political polarization and negative rhetoric (Iyengar and Westwood 2015). RCV advocates claim that because RCV ensures that the victor earns a majority of votes, candidates are incentivized to avoid negative campaigning and moderate their rhetoric to avoid alienating voters (Rank the Vote 2023).

An emerging literature has tested the claim that candidates in RCV elections engage in less negative campaigning than candidates in plurality elections. These studies have found that candidates in RCV elections tend to use less negative campaigning than their plurality counterparts. However, this literature has two key limitations. First, scholars have been unable to demonstrate that the effect of RCV on negative campaigning generalizes across elections in the United States, which are administered according to state law and municipal discretion. Second, past scholarship has not investigated whether differences between candidates and electoral contexts (what I term "electoral features") *condition* the small but significant moderating effect of RCV on negative campaigning. The moderating effect of RCV on candidate rhetoric is small enough that individual variation on contextual variables may enhance its impact on candidates or erase it entirely.

This thesis will address flaws in the literature by investigating whether electoral features condition the use of negative campaigning by candidates in ranked-choice voting elections in the United States. To do so, I leverage quantitative models of campaign rhetoric using sentiment

analysis and statistical methods paired with a qualitative case-study approach that accounts for additional variation and complexity.

The mechanics of ranked-choice voting

In contrast to *plurality* elections, where the candidate who receives the most votes wins regardless of their overall vote share, *ranked-choice voting* elections ensure that the electoral outcomes match the "true" preferences of the electorate. While "ranked-choice voting" refers to several alternative systems of casting and counting votes, almost every district that has adopted RCV in the United States uses *instant runoff voting*, also known as the *alternative vote*. This form of RCV is most appropriate for single-winner elections, and as such is a good model for elections in the U.S., which almost universally has single-member districts (Santucci 2021). Alternative forms of RCV such as single transferable vote or block preferential voting carry their own slate of costs and benefits. Given that the focus of this thesis is U.S. elections, I restrict my analysis to instant runoff elections and treat "RCV" and "instant runoff" as synonymous.

The implementation of ranked-choice voting requires two changes to the process by which votes are cast and then counted. Instead of selecting one candidate, voters may indicate which candidate is their first choice, which is their second choice, and so on, up to a number of choices determined by election administrators (Kimball and Anthony 2018). When all the ballots have been submitted, poll workers use some form of the following procedure to count votes:

1. Allot first-choice votes to their respective candidates.
2. If any candidate receives an outright majority of votes (typically 50% of total votes cast plus one), that candidate is declared the winner and the election is concluded.
3. If no candidate has a majority, eliminate the candidate who received the lowest number (or percentage) of first-choice votes.
4. Allot the second-choice votes of the voters who selected the now-eliminated candidate to their respective candidate choices.
5. Repeat steps 3 and 4 until one candidate has secured a majority of votes. That candidate is declared the winner.

Beyond these general guidelines, the specifics of administering RCV elections vary across electoral jurisdictions, a product of the decentralized nature of electoral administration in the United States. For example, voters in Minneapolis may rank up to three candidates, and more than one candidate may be eliminated each round so long as they all have "no mathematical possibility of winning," a process known as *batch elimination* (City of Minneapolis 2022). By contrast, voters in Alaskan federal and state elections may rank up to four candidates, and only one candidate is eliminated each round (Alaska Division of Elections 2022). Differential implementation of RCV has been treated as trivial by the literature, and there has been no systematic attempt to examine what effects (if any) these administrative differences have on candidate strategy or electoral outcomes.

RCV in the United States: a brief history

Recent implementation of RCV in the United States marks a resurgence of an idea with roots in the late 19th century and the Progressive Era. While forms of ranked-choice voting for use in multi-winner, proportional representation were developed in Europe in the 1850s, the single-winner instant runoff variant was developed by an MIT professor in the 1870s (Ranked Choice Voting Resource Center 2022). Sustained advocacy for the use of RCV began in the 1890s (Santucci and Hughes 2017), when it was part of a broader package of electoral reforms that included ballot initiatives, female suffrage, and open primaries (Barber 2000). The first use of RCV was in 1915, in Ashtabula, Ohio, and over the next two decades, more than two dozen cities adopted the practice for local elections (Ranked Choice Voting Resource Center 2022). However, throughout the 1940s and 1950s, repeal efforts and court challenges were mounted by politicians who lost under RCV. Losing candidates were able to garner public support for their challenges by harnessing broader frustration with machine politics at all levels of government.

Consequently, RCV was eliminated everywhere in the United States it was used except Cambridge, Massachusetts, which has used a variant of RCV (the single transferable vote) continuously since 1941 (Scanlon 2018).

The modern resurgence of RCV in the United States has been led by municipalities, but advocates have also made gains at the federal level. In 2002, voters in San Francisco approved a measure to introduce RCV into local elections, implementing it in 2004 for the first time (San Francisco Department of Elections 2023). Several other cities around the country adopted the practice throughout the 2000s and 2010s. Efforts by RCV advocates reached new heights in 2016, when voters in Maine approved a ballot measure to implement RCV in state and federal elections (Maine State Legislature 2022). These efforts have continued in the intervening time to moderate success. Two states, Maine and Alaska, have implemented RCV for federal offices, including congressional and presidential elections, but no state has carried out a RCV election for state-level offices. According to FairVote, one of the largest RCV advocacy groups, 43 jurisdictions used RCV in 2020. That number increased to 56 different states, counties, and municipalities in the 2022 midterm elections, affecting 11 million voters throughout the country (FairVote 2022).

Negative campaigning

The rise of RCV was predated by the rise of negative campaigning, which first caught the attention of political communications scholars in the 1980s. Debates surrounding the definition of negative campaigning (along with its purpose, predictors, and effects) continue in the literature (Haselmayer 2019). For clarity, I frame negative campaigning in *directional* terms, rather than as an evaluative judgement of individual speech, as articulated by Richard Lau and Gerald Pomper in 2002:

> "Negative campaigning is talking *about the opponent*—his or her programs, accomplishments, qualifications, associates, and so on—with the focus, usually, on the *defects* of these attributes." (Lau and Pomper 2002; emphasis mine)

Negative campaigning may be described as the inverse of "positive campaigning," which is when candidates talk about their past achievements, qualifications, and policy prescriptions. Both methods represent broad strategies of campaign messaging that all candidates employ to some degree. Robust research indicates that negative campaigning may have adverse effects on elections and the electorate, inducing lower turnout and fanning the flames of political polarization (Hansen and Kosiara-Pedersen 2017; Holbrook and Weinschenk 2014; Ansolabehere and Iyengar 1995). Persistent concerns about an increase in the use of negative campaigning, which has only accelerated in the digital age, has prompted a search for solutions.

Ranked-choice voting is one such solution to the problem of increased negative campaigning, and it has been promoted by voting rights activists, academics, and even some elected officials. Advocacy groups like the League of Women Voters and FairVote state that RCV "discourages" negative campaigning (Keller 2017) and that the ability to rank additional candidates "lessens the incentive to run negative campaigns" (FairVote 2023). Candidates at all levels, including sitting U.S. Senators, have endorsed RCV using similar reasoning, suggesting that they perceive the strategic benefits of using less negative campaigning under RCV (Warren and Raskin 2020). Yet no studies have examined what conditions negative campaigning in RCV races from the perspective of candidate strategy, a problem that this thesis will remedy.

Literature Review

Predictors and conditionality of negative campaigning

While negative campaigning in the form of policy criticisms and personal attacks has been employed in U.S. elections essentially since its founding (Callender 1800), the rigorous

study of negative campaigning did not begin until the mid-1990s (Haselmayer 2019). Past research differentiated between positive and negative messages based on whether a message from one candidate was critical of another (Geer 2008). Extensive debates surrounding the qualitative boundaries of what "counts" as negative campaigning continue to plague the literature. Some of this confusion relates to perceived differences between "types" of negative campaigning and their variable effects on voters. Examples of these types and techniques include direct attacks, character smears, policy critiques, contrast advertisements, and party-focused attacks (Lau and Rovner 2009). While these distinctions are useful for developing theoretical models of campaign strategy, most contemporary scholars use a graded approach in conjunction with quantitative methods to analyze messages along the dimensions of sentiment and strength (Rudkowsky et al. 2018). An effective tool for large-*n* analysis of negative campaigning along these dimensions is sentiment analysis, which can identify both the *polarity* and sentimental *intensity* of a text (Liu 2012).

Foundational principles of the negative campaigning literature conceive of candidates as rational actors who use negative campaigning for a strategic purpose, typically to mobilize their own supporters and deter other voters from supporting the opponent (Ansolabehere et al. 1994; Lau 1985). Under this model, candidates "go negative" when the benefits to doing so outweigh the potential risks. Cognitive psychologists have found that individuals devote more attention to negative information and give it more weight (Rozin and Royzman 2001), and political communication research demonstrates that negative messages and advertisements are more likely to be viewed as "newsworthy" and draw the attention of journalists and news coverage (Burden 2002). However, choosing to go negative comes with costs. Excessive or ill-timed negative campaigning can alienate the electorate, causing a "backlash" effect that harms the electoral

performance of the attacker (Walter and van der Eijk 2019), even if the negative campaign is successful in harming opponents (Galasso et al. 2021). Thus, negative campaigning is a conscious and deliberate choice that is subject to the same contextual factors as other types of campaigning: material costs to raise awareness and distribute messages, voter dispositions, and the actions of other candidates.

The Functional Theory of Campaign Discourse developed by William Benoit provides a theoretical context that explains when and how candidates go negative. Functional theory has been applied and validated cross-nationally as a means of analyzing different types of campaign messages, including advertisements (Benoit 2014), televised debates (Benoit et al. 2007), and social media posts (Stein and Benoit 2021). Functional theory views voting as a *comparative* process by which people make judgements about candidates relative to the other candidates in the running (Benoit 2014, 14-16). Under this assumption, candidates distinguish themselves from one another to become a most- (or second- or third- under RCV) preferred choice for voters. According to Benoit, candidates distinguish themselves from their opponents by using campaign messages, which are deliberately structured and delivered to persuade voters (Benoit 2014, 14-16). Functional theory explicitly focuses on candidate strategy and motivation as a driver of message content. When viewed through a functional approach, the purpose of negative campaigning is to make the messenger seem preferable to the target of the message.

The contextual factors that influence when candidates use negative campaigning fall into two categories: candidate attributes and electoral attributes. On the individual, candidate-attribute level, researchers find that incumbents enjoy several advantages over challengers, including increased name recognition, an established record, and specific positive accomplishments (Druckman et al. 2020). These advantages give incumbents strong incentives to run campaigns

centered on positive messaging, which research has shown they tend to do (Valli and Nai 2020). By contrast, challengers must expend extra effort on name recognition, particularly in elections with a crowded field (Druckman et al. 2020), and thus are more likely to rely on negative messaging tactics to receive media coverage and to define themselves as distinct from their opponents (Nai 2020).

Candidate and opponent gender also play a role in campaign messaging strategy. The bulk of this literature examines the influence of societal expectations on how female candidates speak, criticize other candidates, and comport themselves in televised debates and candidate forums (Bauer and Santia 20223; Kahn 2016; Banwart and McKinney 2005, Panagopoulos 2004). The results of these studies are more nuanced than they would first appear. One branch of the literature finds that in addition to talking about different issues, women tend to engage in negative campaigning at a lower rate than men and use different tactics when they do "go negative," such as framing disagreement in terms of policy criticisms rather than direct attacks (Kahn 2016). These outcomes are framed as the result of women internalizing societal expectations that they should act more caring or nurturing than men and the prevalence of biases against women exhibiting traditionally "masculine" traits embodied in directly attacking or criticizing opponents (Rudman et al. 2012). Another branch of the literature finds that under certain conditions, women use *more* negativity in their campaign communications than male candidates (Bauer and Santia 2023). The explanation for this finding: women feel the need to exhibit more "masculine" qualities to be seen as competent leaders and win the votes of particular constituencies (Bauer and Santia 2023).

In contrast to candidate-level variables, there has been less attention to the role that administrative details and electoral systems play in the decision to go negative in American

elections. This gap in the literature is a surprising oversight given the decentralized, state-run nature of American electoral administration. General elections feature more negative campaigning than primary elections, because rival candidates in a primary contest are co-partisans, one of whom will then advance to a general election (Bernhardt and Ghosh 2020). The same principles apply to multi-winner primary elections like "top-two" primaries. Other studies of primary campaigns have found that more candidates leads to a higher rate of negative campaigning, moderated by candidate quality (Peterson and Djupe 2005). The likelihood of victory also plays a major role in determining when to use negative messaging. Candidates, campaign advisors, and political operatives are more likely to advocate for negative campaigning tactics when they are down in the polls (Hassell 2021), which translates into candidates "going negative" when they are behind (Valli and Nai 2020). Even favored candidates become more likely to endorse negative messaging tactics when the race is highly competitive (Hassell 2021).

Effects of negative campaigning

The literature on the effects of negative campaigning has centered on the "demobilization hypothesis" advanced by Ansolabehere and Iyengar in their seminal work 1995 *Going Negative*, which analyzed the effect of negative attack ads on voter turnout. Under this theory, candidates deploy negative advertising to depress turnout for their opponent(s) by making risk-averse voters—including out-partisans and uninformed voters—unsure of who to vote for (Ansolabehere and Iyengar 1995). The actual findings for such a demobilizing effect have been mixed (West 2017; Finkel and Geer 1998), with a contemporary meta-analysis concluding that the effect of negative campaigning on turnout is inconclusive (Lau et al. 1999). Recent studies have even suggested the opposite: negative advertisements inform the decisions of voters by

drawing attention to the weaknesses and harmful policies of candidates (Mattes and Redlawsk 2015).

Additional research on the effect of negative campaigning on voter preferences and electoral outcomes has found mixed results. Negative messaging tactics tend to be effective for challengers to move ahead in the polls, and are less helpful for incumbents seeking to extend their lead (Lau and Pomper 2002). Additionally, studies of races with more than two candidates indicate that there are spillover effects of negative campaigning. While attackers are variably rewarded with more support or a backlash effect and the targets (usually the incumbent) are only mildly and temporarily affected by attacks against them, "idle" candidates not part of the interaction are viewed more favorably and receive more support in controlled studies (Galasso et al. 2021).

Complicating the study of effects on electoral outcomes is the finding that candidates who are behind in the polls use negative campaigning more often. While candidates that use more negative campaigning lose more often, the direction of a causal effect of negative campaigning is difficult to determine given that winning is ultimately decided by vote share and can be predicted by poll position (Lau and Rovner 2009). Despite the findings indicating that negative campaigning is not a guaranteed method to win elections, candidates continue to use negative tactics across different media. Indeed, scholars have noticed an uptick in the use of negative campaigning on social media in recent years, including on Twitter (Hosch-Dayican et al. 2016; Ceron and d'Adda 2016; Gross and Johnson 2016) and Facebook (Baranowski et al. 2022; Auter and Fine 2016).

Negative campaigning and ranked-choice voting

Advocates of ranked-choice voting claim that the system has the potential to reduce (or moderate) negative campaigning and political polarization (FairVote 2023; League of Women

Voters of Maine 2020). They assert that because candidates *must* garner support from a majority of voters to win, candidates will be incentivized to change their strategy. Advocates make the specific claim that candidates will move away from negative campaigning under RCV (Rank the Vote 2023). This claim is likely to drive popular support for RCV, especially among voters dissatisfied by heightened political polarization in the United States(Iyengar and Westwood 2015), but its veracity is a matter of ongoing debate. A small but growing body of literature has sought to test the validity of the "moderating" incentive of RCV in U.S. elections. This work has found evidence for a moderating incentive in some contexts but it effectively assumes that the effect generalizes across the U.S. Such an assumption discounts the possibility that electoral features condition the use of negative campaigning in RCV elections.

Voters in ranked-choice elections perceive that candidates engage in less negative campaigning compared to voters in plurality elections. A 2016 comparative study of ten municipal elections conducted by Todd Donovan, Caroline Tolbert, and Kellen Gracey found that voters in RCV elections reported higher satisfaction with candidate conduct and perceived lower levels of criticism and negative campaigning than voters in a matched set of plurality-election cities (Donovan et al. 2016). These results were consistent with a follow-up study a year later that used both the same cities and a separate sample of Californian cities, constructed using the same comparative strategy of matching RCV and plurality cities (John and Douglas 2017).

RCV candidates attack their opponents and report being attacked less often compared to similar candidates in plurality elections. In a survey of 200 municipal candidates from RCV and non-RCV elections conducted by Douglas and John, 29% of RCV candidates reported being negatively portrayed by their opponents (compared to 40% of non-RCV candidates), while 21% admitted to engaging in negative portrayals, compared to 37% of non-RCV candidates (John and

Douglas 2017). Further research indicates that acting in accordance with the "moderating incentive" of RCV may have downstream effects. A 2021 paper by Courtney Juelich and Joseph Coll using survey data found that when voters perceive that candidates in RCV elections use less negative campaigning, they are more likely to turn out to vote, especially if they are older (Juelich and Coll 2021).

Candidates in RCV elections use less negative campaigning compared to plurality candidates, though conflicting findings indicate that this relationship is complex. Denise Robb conducted a 2011 sentiment analysis of campaign mailers used in elections in San Francisco from 2002-2008 (both before and after the city implemented RCV), and found a sharp drop in negative language and increase in cooperative, positive language after RCV implementation (Robb 2022). A supplemental set of qualitative candidate interviews demonstrated that these changes after RCV implementation were intentional and that candidates thoughtfully considered them (Robb 2022). Martha Kropf performed sentiment analysis on newspaper articles related to elections in ten cities (RCV and non-RCV) and on Twitter posts sourced from candidates in those elections. Kropf found that newspaper articles in RCV cities tended to use more positive than negative words compared to plurality cities (Kropf 2021). Mixed results related to the use of negative campaigning on social media may be attributable to differences between individual platforms and types of messaging. For example, a 2020 working paper that analyzed spending on Facebook ads related to a pair of Maine congressional elections found a shift towards more negative spending once RCV had been implemented. However, the author acknowledges that these findings may not generalize beyond the specific races, which saw an anomalous amount of outside spending on campaign advertising (Clark 2020).

Relatively little work dealing with campaign conduct in RCV races has dealt with direct statements from candidates. To date, only one paper has analyzed the most accessible means of measuring candidate rhetoric—televised debates. A 2020 paper on the subject authored by Eamon McGinn was chiefly concerned with elections in Minneapolis, but also included sentiment analysis of 227 debates from 37 of most populous cities U.S. McGinn found that candidates in debates in RCV cities tended to use more neutral or positive words and phrases than their plurality counterparts, whose statements contained more negative words (McGinn 2020).

While useful for demonstrating the effects of implementing RCV in plurality districts, the McGinn study is emblematic of many of the shortcomings present in every other RCV paper discussed up until this point. First, the paper is comparative, and uses a mixed sample of RCV and plurality elections. While this type of sample can be used to determine *whether* RCV is associated with reduced negative campaigning, the study cannot show *why* that effect occurs or what factors condition it. Second, the pool of elections that McGinn does survey is restrictive, and composed entirely of low-turnout, low-impact municipal elections. Finally, the McGinn study focuses on the effects on voters, sidelining the principal actors who create and distribute campaign messaging—the candidates themselves.

I address the shortcomings present in the literature by examining variation in negative campaigning between RCV elections (i.e., not across RCV and plurality elections) using an expansive sample that improves on past research. This sample is guided by a theoretical model that focuses on candidate strategy, which is the mechanism that determines if and when candidates "go negative" and is affected by changes to the rules of the game, including the moderating incentive of RCV.

Hypotheses

Ranked-choice elections impact candidate strategy in two ways that may have implications for negative campaigning. First, candidates in RCV elections *must* garner a majority of public support in order to win; earning just one more vote than the runner-up is insufficient. If no candidate gets an outright majority of first-choice support, additional rounds of tabulation occur until a majority of voters indicate preference for one candidate. Second, candidates must consider the possibility of getting secondary or tertiary votes from individuals who are not part of their base. A high proportion of voters in laboratory simulation and actual RCV elections tend to rank multiple candidates (Gutiérrez et al. 2022; Coll 2021; Alvarez et al. 2018). Thus, candidates stand to benefit from having broad appeal, especially if they cannot count on winning a majority of first-round support. In fact, candidates often explicitly encourage voters to use their second-choice votes, both in the context of endorsements and as part of broader educational messages about RCV (Armstrong 2019).

That said, candidates are still likely to focus on their base given that secondary support only benefits them under certain circumstances, i.e., if there is no first-round majority. Additionally, voters are likely to hold weaker preferences for the second- and third-choice options compared to the first-choice option (Nielson 2017). Rather than using high-effort outreach methods (e.g., targeted canvassing, policy promises, soliciting endorsements), candidates may attempt to seem generally appealing to second-choice voters. Given the fact that voters tend to be averse to negative campaigning employed by candidates they do not favor, I expect that candidates will use less negative campaigning in RCV races to earn second-choice support and to avoid alienating voters outside of their base.

The formulation of the "moderating incentive" based on the aversion voters have to negative campaigning raises several considerations. Candidates are only likely to seek second-choice support if they believe it will matter and if doing so does not conflict with their first-choice campaigning. A candidate whose first-choice support is dependent on excessive negative campaigning or who believes they will win by outright majority in the first round will not moderate their rhetoric to chase second-choice support.

I analyze negative campaigning as a product of candidate strategy, the proper frame given that the literature portrays negative campaigning as a deliberate and strategic choice. Past studies in the negative campaigning literature have only considered plurality elections. To determine if the conclusions of the negative campaigning literature also apply to RCV elections, I construct the following two hypotheses:

> *H1: Female candidates will engage in less negative campaigning than male candidates in RCV elections.*
>
> *H2: Incumbents will engage in less negative campaigning than challengers in RCV elections.*

While the negative campaigning literature focuses on variation at the candidate level, the ranked-choice voting and electoral systems scholarship considers election-level factors that affect the use of negative campaigning. The ranked-choice voting literature gives special consideration to the number of candidates given that RCV tabulation only occurs if there are more than two candidates in the race. Even if there are one or more clear front-runners in an election, the presence of additional candidates means that secondary support is more likely to be split among multiple candidates.

I extend past research and investigate the effect of variation in how RCV is implemented in the form of *ranking restrictions*, administrative rules that limit the number of candidates voters are allowed to rank. These rules constrain the number of ways that candidates can gain

secondary or tertiary support by allowing them to draw from a fixed number of constituencies. Thus, ranking restrictions reduce the impact of the "moderating incentive" of RCV on candidate strategy. However, ranking restrictions only impact strategy in cases where the number of candidates is larger than the number of ranking restrictions—otherwise, restrictions do not come into effect. Drawing on past findings from the electoral systems literature, I construct the following hypotheses:

> *H3: As the number of candidates in a RCV election increases, the amount of negative campaigning employed by candidates will decrease.*
>
> *H4: In RCV elections with ranking restrictions, the amount of negative campaigning will increase if the number of candidates is greater than the number voters are allowed to rank.*

In addition to the effects *being* an incumbent has on a candidate, the *presence* of an incumbent is a common variable in the negative campaigning and electoral systems literature. Studies in both literatures also stress the importance of *poll position* (at the candidate-level) or *competition* (at the election-level). Competition is especially important for RCV elections because a more competitive election is more likely to proceed to multiple rounds of tabulation. Thus, the incentive to moderate and use less negative campaigning should be stronger in more competitive elections. Candidates have more to lose if they alienate voters by engaging in negative campaigning. Poll position as a component of candidate strategy prior to Election Day is difficult to measure for municipal elections given that few local races have robust and regular pre-election polling. Thus, I gauge the interaction of the competitiveness of a district with the prevalence of negative campaigning in an election. Both literatures contribute to my last set of hypotheses:

> *H5: Candidates in open-seat races will engage in less negative campaigning than challengers in races where an incumbent is present.*

H6: Candidates in more competitive races will engage in less negative campaigning compared to candidates in less competitive races.

My six main hypotheses are informed by past RCV and negative campaigning research, but I also consider alternative possibilities. Candidates may not perceive the pressures inherent to an RCV election, or their strategy may discount the importance of additional rounds of voting and favor an attempt to win an outright majority. If this is the case, I would expect to see candidates employing similar levels of negative campaigning across both candidate- and election-level conditions, with no significant variation between elections. Polarization in the electorate may have intensified to such a degree in recent years (Iyengar and Westwood 2015) that going negative is a superior strategy to reaching a broader group of voters. If this is the case, polarization and playing to the base may swamp the moderating incentive of RCV in terms of candidate strategy, given that the moderating incentive induces modest effects on behavior. This possibility is especially salient in an all-RCV sample, as most RCV elections are held in large urban centers that reliably lean Democratic and are likely to be polarized in a left-leaning direction. In this case, I would expect to see increases in negative campaigning associated with more recent elections compared to older elections, and an increase in negative campaigning in primary elections, which are typically lower-turnout and contain a larger percentage of committed partisan voters (Hirano et al. 2010).

Methods

While past research has investigated the effect of RCV on rhetorical moderation and negative by comparing plurality and RCV elections, I take an additional step and measure the *conditionality* of that effect. To do so, I make comparisons between as many RCV elections as possible using fine-grained measures of negative campaigning. Toward that end, I have assembled the largest collection of RCV elections in the United States to date. I complement the

statistical analyses of those data with qualitative case studies of three federal elections in Maine. The case studies illustrate the complex interaction of my independent variables and other contextual factors with the use of negative campaigning.

Election selection: constructing a dataset of RCV races

The first major contribution of this thesis to the literature is the dataset that I use to test my hypotheses, which is more comprehensive than any previous attempt to gather information about RCV elections in the United States. I collect data on 146 unique candidates who ran in 34 different RCV elections between 2009 and 2022. These data help to paint a more comprehensive picture of the current state of RCV and ensure that there is sufficient variance on the independent variables for hypothesis testing.

The elections in my dataset consist of all 21st-century RCV federal elections and mayoral elections in major U.S. cities, which I identify using the RCV legislative implementation tracker from FairVote, an electoral reform advocacy organization (FairVote 2022). Negative campaigning likely pervaded the 20th-century municipal elections that used RCV, but relevant data that serves as an indicator of candidate strategy (e.g., transcripts of debates) are unavailable. I source data about the candidates (gender, incumbent status) and the election (number of candidates, competitiveness, ranking restrictions, presence of an incumbent) from Ballotpedia, supplementing with the websites of various Secretaries of State and electoral returns. I elaborate on individual cases and provide descriptive information in Chapter 1.

Negative campaigning data

My dataset is comprehensive both in terms of how many elections it contains and *how* it captures negativity. To determine the amount of negative campaigning, I use a mixed approach that incorporates sentiment analysis of candidate Twitter posts, televised candidate debates, and

television advertisements. I provide an in-depth discussion of these methods in the proceeding chapters.

In recent years, it has become increasingly common for candidates to campaign on Twitter, and tweets have been validated as a source of campaign sentiment data (Gross and Johnson 2016; Bekafigo and Pingley 2015). Negative communication receives higher engagement than positive communication on Twitter, thus creating an incentive for candidates to skew negative. However, additional research shows that these incentives do not meaningfully impact total campaign sentiment over an election (Mueller and Saeltzer 2022). Following Kropf (2021), I use the Twitter API and various web scraping tools to collect tweets during the "peak campaign season" in the sixty days prior to Election Day (Kropf 2021). In the case of primary elections, I collect all tweets in the sixty days prior to the date of the primary. Finally, I engage in a small amount of text preprocessing, removing Twitter-specific characters like "@" and "#" to allow for proper interpretation by sentiment analysis models.

Candidate debates, whether broadcast on television or hosted by a community group, are a unique example of campaign messaging that make them ideal for the study of candidate strategy across different levels of election. Outside of minor rule changes regarding speaking time and opportunities for rebuttal, the format of candidate debates remains largely consistent across electoral districts, race level, and type of office. Even local-level debates are held with an audience in mind, and those that are broadcast appear on trustworthy news channels during high-viewership timeslots. Though candidates may rehearse specific messages, phrases, and sound bites before a debate, debates also contain many instances where candidates are forced to speak off the cuff or respond in the moment to attacks from their opponents. These and other factors

ensure that debates remain an important source of data for understanding candidate messaging strategy and tracking campaign sentiment.

To find footage of candidate debates to use in sentiment analysis, I conduct a manual search across different web video hosting platforms (e.g., C-SPAN, YouTube) using a specific set of keywords, which I supplement with publicly available media archives of local and national news organizations. I then converted the audio to text using the highly accurate Text-to-Speech tool provided by Google Cloud (Siegert et al. 2020; Xu et al. 2020; Wang 2021), supplementing the machine learning model by manually adding speaker diarization to match statements to speakers. Each "statement" in my set of debate data consisted of a single sentence spoken by an individual candidate, or a sentence fragment in cases where multiple candidates spoke at once or cut each other off.

Past studies of negative campaigning in both RCV and plurality elections have often examined the content of advertisements, which provide a good indication of candidate strategy. Advertisements are public-facing and are designed with wide viewership in mind, and their high cost means that candidates must be strategic about when and how to use them. Ads are also a key vehicle for negative campaigning and have measurable impacts on voter behavior (Ansolabehere and Iyengar 1995). However, advertisements are by their nature artificial, and are not as direct a measure of candidate sentiment as debates or tweets. Furthermore, most reliable databases of political advertising data only track federal elections, whereas the vast majority of RCV elections in the U.S. have been local or municipal. Past studies of RCV elections by political communication scholars have been able to make use of advertising data because their study was restricted to a handful of elections, using methods that do not scale for my data.

I use advertisements as a supplement to my main set of campaign messages, using data from the Wesleyan Media Project, a group of researchers who have analyzed political broadcast television advertisements from the Campaign Media Advertising Group since 1996.[1] These data include information about ad cost, distribution, and viewership, an assessment of ad tone and sentiment, and various associated metadata, a full list of which is available in the codebooks in the online supplementary material.[2] The data only exist for the federal elections in my sample, so I incorporate a limited analysis of ad sentiment in Chapter 3, separate from my main results.

Sentiment analysis

Rather than categorize the statements in my dataset as instances of "positive" or "negative" campaigning, I use two dictionary-based sentiment models to provide nuanced negativity ratings (Liu 2012). I employ the Valence Aware Dictionary for sEntiment Reasoning (VADER) and supplement my results with the 2015 Linguistic Inquiry and Word Count, otherwise known as LIWC (Pennebaker et al. 2015; Hutto and Gilbert 2014). I discuss how VADER and LIWC rate sentiment in additional detail in Chapter 2. For each statement, I collect the compound sentiment score provided by VADER, then aggregate the statement-level results for each candidate and each election. The resulting value is an indication of how negative statements made by a candidate or in an election were, operationalized between -1 (most negative) and 1 (most positive). I then use a multilevel statistical model to determine the relationship between my independent variables and the average amount of campaign negativity for each candidate and election as calculated by VADER. This model indicates which candidate-

[1] Both datasets available thanks to a generous grant from the Nelson A. Rockefeller Center for Public Policy and the Social Sciences at Dartmouth College.
[2] All online supplemental material, including additional results, replication data, and reproducible code, is publicly available and stored on the Harvard Dataverse: https://doi.org/10.7910/DVN/XW95OP

and election-level variables condition the use of negative campaigning by candidates in different RCV elections.

In Chapter 3, I employ a qualitative case-study method to study additional aspects of RCV elections that are difficult to quantify or measure for all elections in my main sample. I gather additional data on a restricted set of elections for which those data are available: federal RCV elections since 2018. I integrate these new data, which includes campaign finance filings and polling results, with the results of my sentiment analyses and conduct additional background research onto the candidates and context of each election.

Findings

My findings indicate that the moderating effect of ranked-choice voting on negative campaigning is conditional on candidate-level factors like gender and incumbent status, as well as the presence of an incumbent in the race. These results support H1, H2, and H5. Female candidates are 10-14 percentage points less negative (i.e., more positive) than male candidates, while incumbents are 14 percentage points less negative than challengers. Races with an incumbent are 10-12 percentage points more negative than open-seat races. Additionally, I do not find evidence of an interaction between gender and incumbent status regarding negative campaigning. These findings match the conclusions drawn from the negative campaigning literature about candidate-level variables and serve as an indication that other findings from plurality-based research may also apply to RCV elections.

I lack support for my other hypotheses related to election-level variables, including the number of candidates, ranking restrictions, and competitiveness. The lack of a distinct relationship between the independent variables and the use of negative campaigning by candidates is an indication that the moderating effect of RCV may be robust (i.e., present and

similarly effective) across different electoral contexts. Differences in how RCV is implemented by states and municipalities do not seem to impact how and when candidates use negative campaigning in RCV races.

The results of my qualitative case studies indicate that the effects I find may be more complex than they initially appear. While I find evidence that candidate-level variables condition the use of negative campaigning *on average*, analyzing even a handful of elections in detail indicate that the real-world manifestation of the moderating incentive of RCV is more complex. My research indicates that candidate- and election-level variables may be stronger in different contexts, particularly when the outcome of an election is uncertain, often because the election is highly competitive or perceived as such. However, I also argue that the moderating incentive of RCV may be critical to the outcome of an election under certain circumstances. I also illustrate one case where I believe the choice of one candidate to reduce their negative campaigning led to their success. While these qualitative results are preliminary, they indicate the need for further qualitative analyses of campaign strategy, particularly those that make use of interviews or interaction with candidates and their campaign staff.

Roadmap

The conditionality of the moderating effect of RCV on the decision to "go negative" is the central focus of this thesis, but such a complex question cannot be answered all at once. I break my investigation and its novel contributions into multiple parts, as follows. In Chapter 1, I discuss my novel dataset of campaign messaging in RCV elections, which encompasses more than 38,000 statements sourced from 34 elections, and present relevant descriptive statistics. I reserve the bulk of my statistical analyses for Chapter 2, including the results of my sentiment analyses and my multilevel statistical model, which measures the relationship between my key

independent variables and campaign negativity. Chapter 3 incorporates advertising data into a restricted set of elections and contains three qualitative case studies, which put the moderating effect of RCV into context with other considerations that candidates make throughout the campaign. I conclude with a review of my findings, a discussion of the limitations of this study, and guidance for future research.

Chapter 1: A new ranked-choice voting dataset

The use of RCV in U.S. elections has expanded rapidly over the last two decades and has accelerated since Maine adopted RCV in 2018 for federal elections. The scholarship surrounding electoral reforms (in general) and RCV (in particular) has not kept pace with these developments outside of investigating the effects of RCV on turnout (McCannon 2022; Coll 2021; McDaniel 2019). This gap in the literature exists in part because RCV elections have largely been municipal, which are difficult to analyze with conventional methods due administrative variability and a lack of available data.

This thesis represents the first attempt to assemble a comprehensive dataset of candidate rhetoric and messaging strategy in an exclusively RCV sample within the United States. In this chapter, I lay out the process by which I collected the data, describe how I use them to measure campaign negativity, and explain how I evaluated the electoral features of each race. I then present descriptive statistics about the dataset. I conclude by discussing how I structure these data for my analyses in Chapter 2, in which I test the conditionality of the moderating incentive of RCV on several electoral features.

The dataset encompasses information about 168 candidates who ran in 34 local and federal elections between 2009 and 2022. I count individuals who ran in multiple elections as separate candidates, as candidate- and election-level variables (including incumbent status!) may vary between election cycles. My sample of 168 candidates thus consists of 146 unique individuals. I sampled 24,373 candidate tweets and 12,607 statements made during televised candidate debates, the latter of which are separated at the sentence level. I find sufficient variance on my selected independent variables at both the candidate-level (gender, incumbent status) and the election-level (competitiveness, incumbent presence, number of candidates,

ranking restrictions) to draw robust conclusions about the relationship between these variables and the negativity of campaign messages.

Data Collection

Case selection

I examine a set of elections that is much larger than most studies of RCV, which often compare limited set of RCV and non-RCV municipalities (Kropf 2021; McDaniel 2018; Neely and Cook 2008). I collect data for all RCV elections conducted in the United States between the years 2009 and 2022 at the mayoral level or above. However, my dataset is not *fully* comprehensive. It excludes two types of rarely-studied elections: races for local executive offices, such as selectboard member, in small towns (e.g. Telluride, CO, population 2,595) and elections for municipal positions in larger cities (e.g. San Francisco City Attorney).

I exclude local elections below the mayoral level due to data availability. Gathering information on mayoral elections for large cities proved to be a complex, labor-intensive process that involved social media scraping and transcribing archived videos of candidate debates. For races below the mayoral level in large cities, or races at any level in smaller towns, the relevant data do not exist, or would require a local expert to obtain and interpret. Furthermore, these elections have lower levels of turnout than mayoral elections (Caren 2007), which themselves have low turnout in comparison to federal races (Hajnal and Lewis 2003). Thus, I restrict my sample to local mayoral elections and federal elections in Maine (Senate, House Districts 1 & 2), and Alaska (at-large House District). Prior studies of ranked-choice voting have demonstrated that mayoral elections attract enough attention, turnout, and money to reliably analyze civility and negative campaigning using methods designed for use with federal elections (Kropf 2021;

Donovan et al. 2016). A full list of the elections in my dataset is available in the online supplementary materials.

Independent variable data: candidates and elections

I source the data for my independent variables from Ballotpedia, a politically independent, online nonprofit that collects information about elections and ballot referenda at all levels in the United States (Simon and Tatalovich 2022; Gastil 2022). I supplement these data with information from state and municipal websites and the offices of individual Secretaries of State. At the candidate level, I obtain information regarding the officially declared victor, a full list of registered and write-in candidates, candidate gender, and incumbent status. At the election level, I collect information on the presence of an incumbent, number of rounds of RCV tabulation (if any), final margin of victory, and ranking restrictions. Whenever possible, I verify this information—particularly the list of candidates, victor, and margin of victory—by referencing publicly-available

My last independent variable is the competitiveness of the election, which was difficult to measure given that I am considered with how competitive candidates *perceived* the race to be, which cannot be determined by post-election measures like vote share. Most races in the sample take place in large municipalities, which lean Democratic and have sparse data compared to federal- or state-level elections. For my competitiveness indicator, I use historical data for all races, even when third-party or scholarly estimates of partisan competition exist. For each election, I take the average of the margin of victory in the two previous election cycles. This approach is a parsimonious and consistent method for single-member districts under both plurality and instant runoff electoral systems (Cox et al. 2020). Analyzing RCV elections under parliamentary systems or proportional representation rules may require different methods

(Cronert and Nyman 2021), but these methods are applicable for U.S. elections for single-member districts. In cases where one or both previous cycles was an RCV election and more than one round of vote-counting occurred, I follow recent scholarship and capture the *final* margin of victory. Research indicates that use of the final margin of victory is appropriate both in studies that make comparisons between RCV and non-RCV elections and multiple RCV elections (McDaniel 2019).

Modeling electoral competitiveness is a complex task, and my chosen method is not without its flaws. Using historical competitiveness is a good method for determining how competitive elections were in the past—in a sense approximating the prior beliefs candidates hold about the race—but it does not consider changes between electoral cycles. Even a two-year gap between elections may contain events and factors that alter the state of the race. Candidates may be caught up in scandals, the demographics of a district may change, and former electoral heavyweights may die or drop out of the running. All these events and more may also take place *during* the campaign, further complicating a competitiveness measure in cases where regular and robust polling data tracked the views of the electorate. I attempt to correct for outlier events that might impact the competitiveness of a race, including the departure or death of an incumbent, by using an averaged two-cycle indicator, "smoothing out" the impact of any given event. Additionally, I acknowledge that competitiveness is best measured as a sliding scale that changes based on the contemporaneous events of an election. However, gathering data at a sufficient level of detail for all elections in my sample is impractical and, given the availability of even basic data, likely impossible. To approximate candidate *perceptions* of competitiveness, I retain my historical competitiveness measure, even though it may be an imperfect indicator of how competitive the race seemed at its inception.

An alternative but inferior approach to measure competitiveness involves the use of post-election measures, such as total vote percentages. While outcome metrics may be used to determine if Election Day was competitive, they provide no information about how competitive the race was at any other point during the campaign. In other words, post-election measurements of competition do not give any insight into how competitive the candidates *perceived* the race to be. As such, total vote share cannot be used to evaluate candidate strategy *during* the campaign.

Dependent variable data: tweets and debate statements

To measure the level of negativity embedded in campaign messaging, I use candidate tweets and statements made during televised debates, two direct sources of campaign sentiment that have precedence in the literature. The constraints imposed on tweets (character limits[3]) and debate statements (speaking time limits) incentivize candidates to communicate strong sentiments in a clear manner that audiences will understand (Mueller and Saeltzer 2022; Walter and Vliegenthart 2010). While candidates may outsource management of their Twitter accounts to staff (Parmelee and Bichard 2011, 8-9) and sometimes deliver prepared statements in debates, both types of data represent the actual words that candidates say in public. In the digital age, candidates can receive near-instant feedback from the public on their Twitter messaging and debate performance. Indeed, candidates with sufficient staff may collect data both after and *during* debates by scouring social media, allowing them to adjust their messaging strategy (Lee and Xu 2018; Kalsnes et al. 2014; Mascaro and Goggins 2012). Finally, candidates can and do respond to one another in real-time in debates and on Twitter. The potential for dialogue allows for analyses of candidate messages in isolation and alongside one another (Hemsley et al. 2018;

[3] In February 2023, Twitter raised the character limit on tweets to 4,000 (Reimann 2023). All tweets in this sample were created 2009-2022 under either the 140- or 280-character limit.

Maier and Jansen 2017). Table 1.1 shows key features of the data collection process for Twitter and debate data, including time frame, search protocols, and relevant exclusion criteria.

Table 1.1: Statement search and text processing protocols

Medium	Timeframe	Search method	Processing	Total candidates	Total statements	Relevant exclusion(s)
Twitter	Sixty days prior to Election Day through Election Day	Manual search (account) API Request + web scraping (tweets)	Scrape statements, text preprocessing	93	26,353	Account created prior to election Account verified **or** attributable based on biographic information Account is public Account has at least 1 tweet
Debates	Any time prior to Election Day	Manual search (C-SPAN, YouTube, local news archives, etc.)	Google Cloud Text-to-Speech, manual speaker diarization, separate sentences, text preprocessing	118	12,607	Full debates Debates with 2+ candidates 1 debate per election

Note: Candidates who do not have any statements according to these protocols are dropped from the sample. Elections for which no candidates have statements according to these protocols are also dropped.

Tweets

Given the wide variation of candidates visibility and social media presence between elections, I collected Twitter accounts by manual search. I initially searched by name (first and last) and location, retaining only those accounts that could clearly be attributed to a candidate or their campaign and which were created before Election Day. For example, in the case of Oakland mayoral candidate Libby Schaaf, I collected data from both @libbyschaaf and @libby4mayor. In some cases, finding the correct account(s) required more scrutiny. For example, Maine state senator Jill Duson has three public accounts: @JillDuson, @Duson4Portland, and @Jill4Senate. @JillDuson was created in 2010 and can be used as a source of data for the 2011 mayoral

election in Portland. However, @Duson4Portland was created in 2016, and thus is excluded from the 2011 count. Likewise, @Jill4Senate was created in 2022 and was used as a campaign account when Duson ran for state legislature, so all tweets from that account are excluded as well.

I used R to interface with the Twitter API and collect tweets from the set of candidate accounts, supplementing with web scraping tools developed by Octoparse and ScrapeHero for accounts with more than 3200 total tweets.[4] Following Kropf (2021), I collect tweets during the sixty days prior to Election Day, a period that has been dubbed the "peak campaign season." In the case of primary elections, I collect all tweets published during the sixty days prior to the date of the primary. For candidates who competed in both an RCV primary and RCV general election, I collected all tweets published during the sixty days before the primary through the date of the general election. Finally, I engage in a small amount of text preprocessing, removing unnecessary instances of "@", "#", non-ASCII characters, and "RT," which indicates a retweet. Text processing of this type is necessary to prevent sentiment analysis tools from being confused by strings that appear like words but in fact contain no relevant information.

I do *not* collect tweets published between the end of Election Day and the declaration of a winner. This "waiting period" can be longer in RCV elections, as demonstrated by the 2021 NYC mayoral election (Matthews and Hajela 2021). Scholars have raised concerns about the rhetoric politicians engage in during the vote-counting process, specifically that candidates may attempt to interfere with the process of tabulation and overturn the election (Saedeen and AlBzour 2022; Salama 2022; Adelman 2021). While important, this body of work is inapplicable to the present study, which is concerned with attempts by candidates to persuade citizens and

[4] At time of data collection, Twitter's API access for undergraduate researchers only allowed requests of the 3,200 most recent tweets from any given account. Scraping was not necessary for some accounts associated with recent elections that had more than 3,200 total tweets.

garner support in the form of votes. No matter how long the post-Election Day tabulation takes, no more votes may be submitted after the polls close, making candidate messaging strategies irrelevant to the outcome. Concern about the use of social media to contest the results of an election is warranted, especially as it pertains to outbreaks of political violence like the January 6, 2021 attack on the U.S. Capitol (Kennedy et al. 2022; Fuchs 2021). However, such radical attempts to subvert electoral results are not present in my dataset and lie outside the scope of this thesis. I discuss the only prominent case of a candidate engaging in legal action to affect the outcome of an RCV election in the United States in Chapter 3.

Debate statements

To find footage of candidate debates for sentiment analysis, I conduct a manual search across different hosting platforms a specific bank of keywords. I searched for "[Race Name]" + "[Race Year]" + "[debate/candidate debate/candidate forum]" across C-SPAN, YouTube, the Vanderbilt Television News Archive, the Internet Archive, and Google. For example, to find candidate debates for Alaska in 2022, I searched "Alaska special election congressional at-large district 2022 debate" across the above platforms. I include various iterations that contained some or all of the above words. In cases where more than one debate was available, I selected the debate closest to the election to avoid bias towards newer cases (easier to find) and higher-level elections (more available debates). In some instances, I only found a print, radio, or television news piece *about* the debate, but not the debate itself. For these cases, I consulted whatever publicly available archives existed for the relevant media outlet in order to find a recording of the debate in question. Taken together, these methods were successful; I found 31 debates, one for almost every election.

To prepare the debate statements for text analysis, I downloaded the video or audio recordings of the individual debates and applied the Google Cloud Text-to-Speech tool, which has been validated for high-quality audio and speech processing (Wang 2021; Siegert et al. 2020; Xu et al. 2020). On average, the transcriptions tool performed well, with some variability attributable to the quality of the original recording (e.g., raw camera feeds, recordings of Zoom meetings). The transcription tool also performed poorly when multiple candidates were speaking at once. These moments of "crosstalk" were frequent enough that I manually added *speaker diarization* to each transcript. For each debate, I separated statements by speaker identity while also verifying the accuracy of the Speech-to-Text tool at random intervals. As with the Twitter data, I engaged in light text preprocessing (removing moderator statements, standardizing name spellings, etc.) before beginning sentiment analysis.

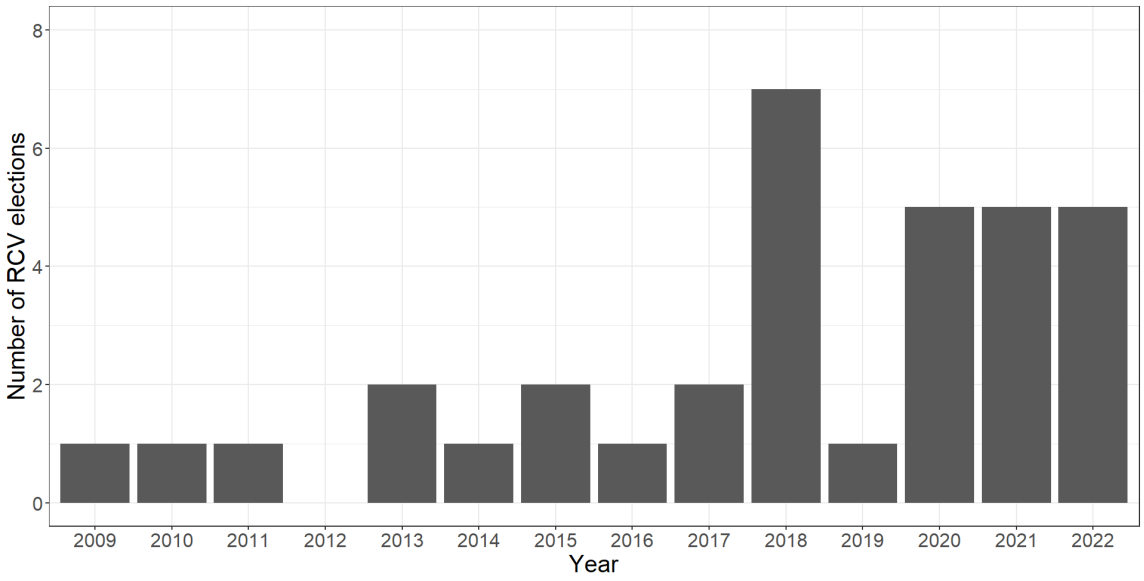
Descriptive Statistics

Cases and data availability

The exclusion criteria detailed in Table 1.1 leaves 168 candidates (146 unique individuals) with at least one tweet or debate statement in the dataset. The 34 elections in the sample were contests for 16 unique offices—the "early adopters" of RCV administered multiple elections during the studied time frame. Only four elections were primary elections. The distribution of eligible elections displays some bias towards more recent elections, as demonstrated by Figure 1.1. This "recency bias" is the result of cumulative adoption of RCV over time. More states and municipalities have begun using RCV in the last two decades than have stopped using it, and all municipalities that implemented RCV have conducted at least one RCV election. The spikes in 2018 and 2020 are driven by Maine adopting RCV and reporting on ensuing legal controversies in state and federal courts. Many of the municipalities in the sample

only hold an election for municipal offices (e.g., mayoral election) every three to four years. More recent RCV adopters have held only one or two RCV elections. Maine has held eight RCV elections since it implemented RCV in 2018, a consequence of having two congressional districts and the schedule for its Senate seat elections.

Figure 1.1: Distribution of RCV elections by year



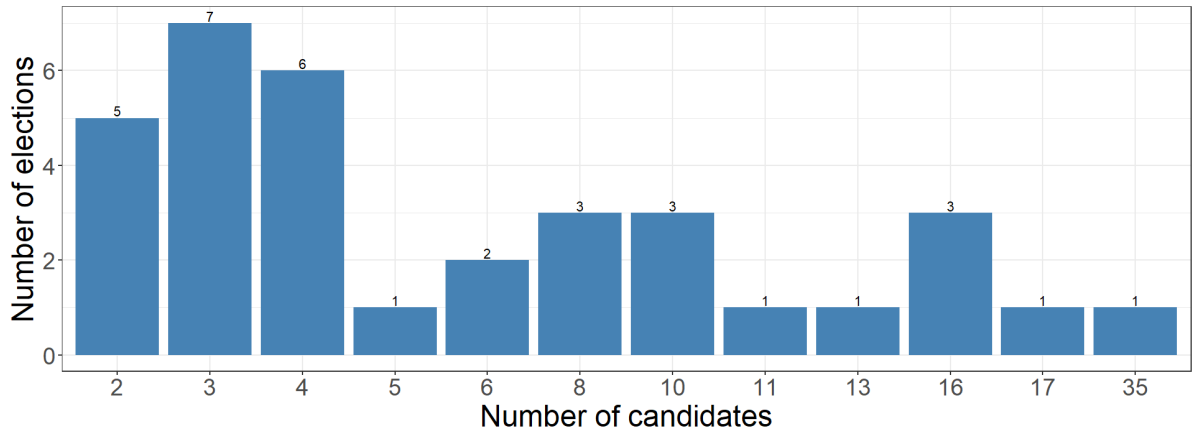
Note: Number of elections reflects the total number of RCV elections at the mayoral level or higher each year.

A count of the number of RCV *elections* per year is the proper tool for gauging RCV adoption and implementation, but its lack of correlation with the number of *candidates* per year may cause confusion. For example, 39 candidates ran in RCV races in 2013 despite that year only having two RCV elections. This was because of the withdrawal of a longtime incumbent from one of the races (Minneapolis mayoral), which stimulated candidate entry and resulted in a final candidate pool of 35 candidates. The multilevel modeling method I use in Chapter 2 accounts for potential issues raised by having distributions of candidates that do not uniformly match the distribution of elections across the years.

Independent variables

Figures 1.2 and 1.3 report the distribution of the number of candidates and the maximum number of candidates that voters could rank. The mean number of candidates in an RCV election was 8, with a median of 5. These are larger fields than most federal U.S. elections, though whether RCV induces more candidates to enter elections remains a topic of scholarly debate (Santucci and Scott 2021). Many of the sampled elections take place in urban centers which lean to the political left (Parker et al. 2018), causing the elections to resemble Democratic primaries instead of multi-partisan contests. The fact that elections operating under quasi-primary conditions had more candidates fits with research showing that primary elections, whether single- or multiparty, have larger candidate pools than general elections (Hirano et al. 2019, 2). Additionally, local elections have lower barriers of entry than the state- or federal-level elections, both in terms of legal requirements, e.g., filing and registration costs, and organizational requirements, e.g., voting population (Schuerman 2010; Schleicher 2007).

Figure 1.2: Number of candidates in sampled elections

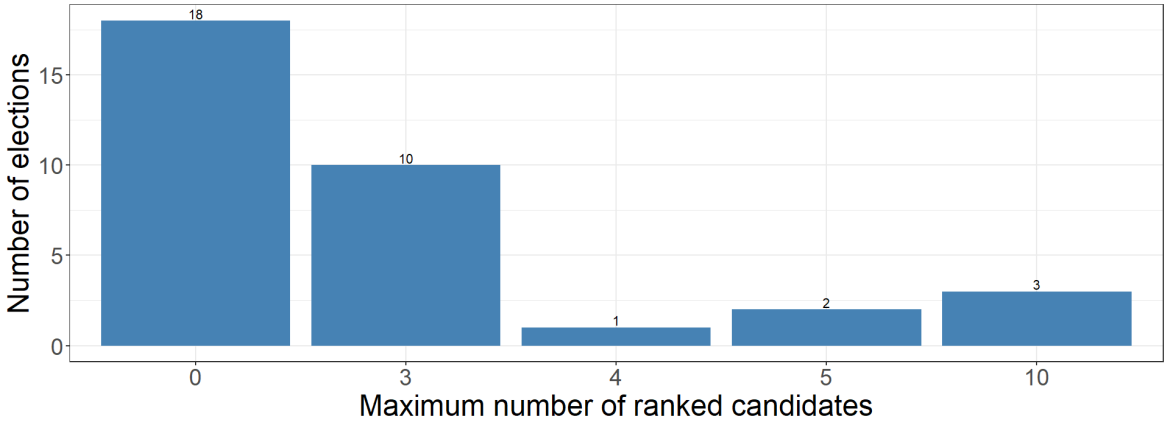


Note: Bars represent the number of elections in which that number of candidates were on the ballot. Elections with only two candidates are retained despite functioning as plurality elections.

Figure 1.3 shows the level of variation in statutory limits on the number of candidates that could be ranked, from 0 (no limit) to 10. The "No limit" category refers to states or

municipalities that allowed voters to select as many candidates as they preferred or that instituted RCV without imposing a limit on the number of candidates voters could rank. Given that 18 elections (53%) had no restrictions and the limited variation in races with restrictions in other numerical categories, I dichotomize the ranking restrictions variable. In the analyses reported later, an election is scored "1" if there were restrictions on the number of candidates that could be ranked and "0" if there were no restrictions.

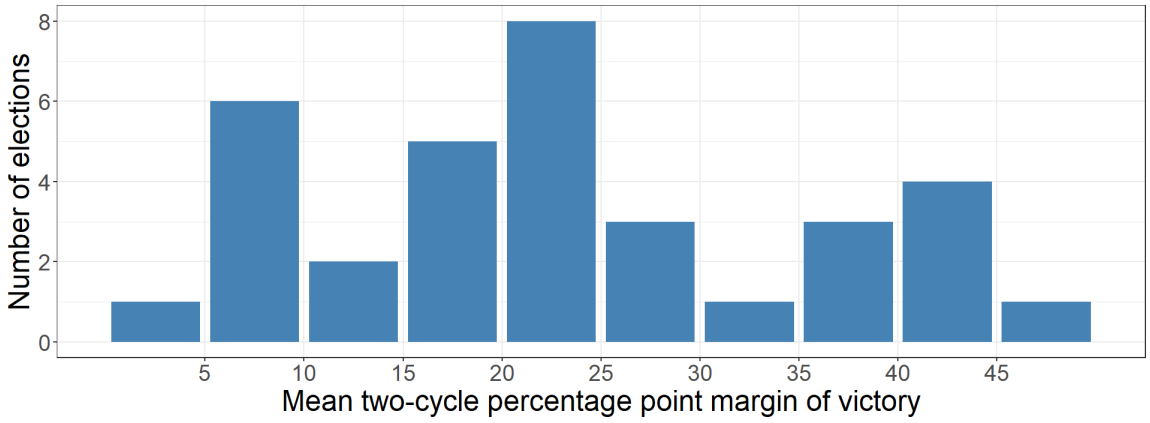
Figure 1.3: Distribution of ranking restrictions in RCV elections



Note: In races where the "maximum" number of ranked candidates was 0, there was no limit to the number of candidates that voters were allowed to rank, i.e. ranking restrictions were not present.

Figure 1.4 summarizes the distribution of the competitiveness measure, which is represented as the mean margin of victory across the two previous election cycles. There can be substantial variation in margin of victory depending on the presence of individual candidates, especially longtime incumbents. The two-cycle measure works to smooth excess variation in cases where longtime incumbents ran in one cycle but not the following cycle, allowing the measure to be more representative of the competitiveness of the *district*. As shown in Figure 1.4, there is wide variation across the competitiveness variable. The mean two-cycle historical competitiveness across the elections in the sample was 24.7 percentage points, with a median of 23.1 percentage points and a standard deviation of 12.9 percentage points.

Figure 1.4: Historical competitiveness of RCV elections



Note: Margin of victory rounded to the nearest integer and binned.

Some cases are excluded from Figure 1.4. I exclude three primary elections in which the victor ran unopposed in the two previous elections, which would yield a margin of victory of 100%. Additionally, only one cycle of historical competitiveness data is available for the 2015 mayoral election in Portland, Maine. Prior to 2011, Portland had not held a citywide election for mayor since 1923 (Maine Voices 2010). For these four cases, I impute a value for historical competitiveness, which is represented by a random value within one standard deviation of the mean level of competitiveness across the dataset.

Table 1.2 summarizes information on additional independent variables. Only 50 of the candidates in the sample (29.8%) are women. However, as mentioned previously, the significance of this distribution is difficult to interpret given that the sample spans twelve unique election years between 2009 and 2022. Twenty races, or 58.8% of sampled elections, had an incumbent, several of whom ran multiple times. However, the large size of many RCV candidate pools and the fact that all races in the sample are for single-member districts means that the proportion of incumbents in the total candidate pool is relatively small. There are twenty incumbents in the sample. As mentioned previously, most elections are unrestricted, and allowed

voters to rank any number of candidates, up to the total number of candidates in the election plus one to allow for write-in candidates.

Table 1.2: Distribution of additional candidate-level and election-level attributes

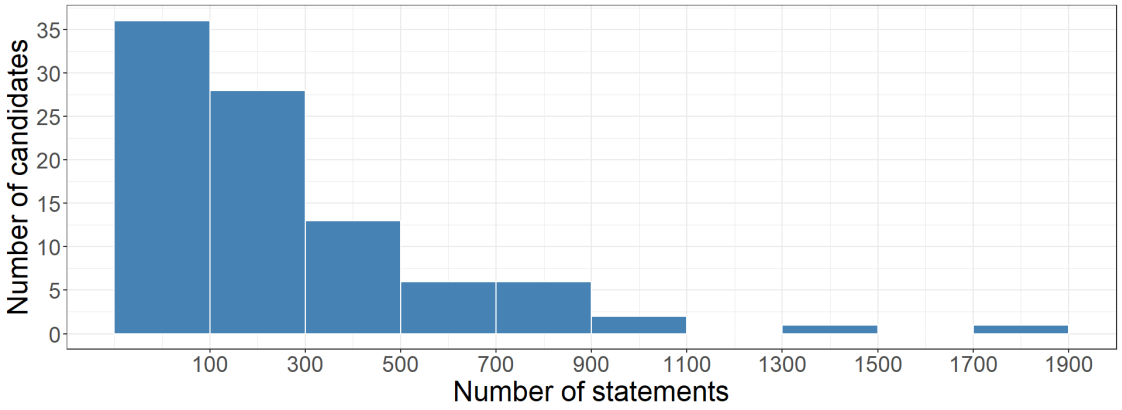
Variable	**Level**	**Percentage of cases**
Gender	Candidate	29.8% female 71.2% male
Incumbent Status	Candidate	11.9% incumbent 88.1% challenger
Incumbent Presence	Election	58.8% incumbency 41.2% open-seat
Ranking Restrictions	Election	47% have restrictions 53% unrestricted

Note: Percentages calculated separately for candidate- and election-level variables. Candidate-level variables based on total number of candidates, including those who ran multiple times in different elections.

Tweets

A majority (71%) of the sample of 168 candidates published at least one tweet in the sixty-day period leading up to Election Day. Among this subsample of candidates (the "active Twitter users"), the mean number of tweets posted was 257, with a median of 133 tweets. Comparable data regarding candidate Twitter use serves as a useful point of comparison despite only being available for federal legislators. The median candidate in my dataset tweets more frequently than members of Congress in 2016 (66 tweets per month compared to 59 tweets per month), but less frequently than members of Congress in 2020, who published 107 tweets per month on average (Van Kessel et al. 2020). Because these data span a thirteen-year period, it is difficult to draw conclusions about any given candidate from these comparisons. Figure 1.5 displays the full distribution of Twitter usage by active Twitter users in the sample. The present dataset covers nearly the entire time Twitter has existed, including a crucial period (2015-2020) during which its usage in political campaigns became mainstream (Van Kessel et t al. 2020). Increased usage of Twitter for campaigning in more recent years among candidates in my sample contributes to the slight recency bias in my data (Van Kessel et al. 2020).

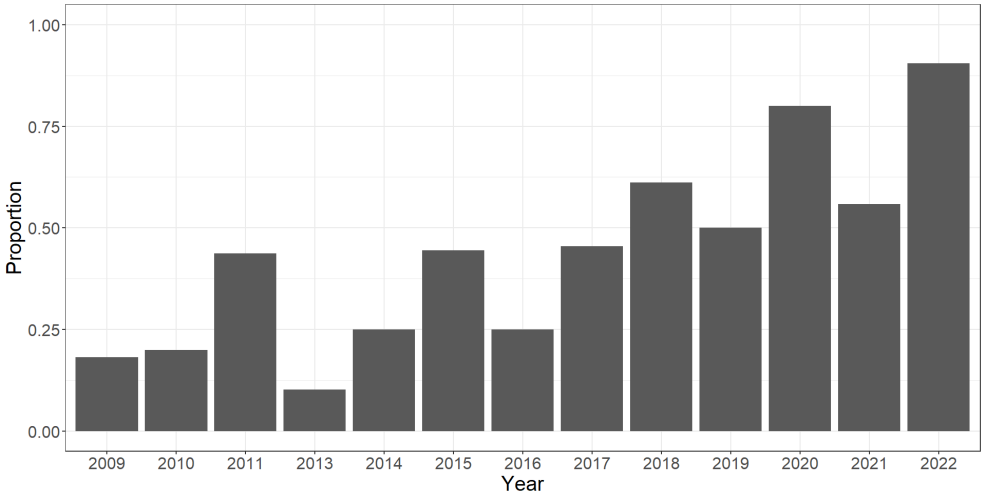
Figure 1.5: Distribution of tweets among RCV candidates



Note: Candidates who did not tweet excluded from distribution. Number of tweets includes quote tweets (retweets with additional text added by candidate). Tweets gathered over a period beginning sixty days before Election Day.

As Figure 1.6 demonstrates, more than half of candidates in the sampled elections from 2018 onward used Twitter, compared to much lower percentages from 2009-2017.

Figure 1.6: Proportion of RCV candidates that use Twitter per year



Note: Proportion of candidates calculated as number of candidates in sample that used Twitter divided by total number of candidates in the sample per year.

Additional factors that contribute to recency bias in my Twitter sample include a marked increase in the use of social media during the COVID-19 pandemic and data availability, which has been studied in political science and across other disciplines (Cellini et al. 2020; Humphry 2020; Maizland 2020). Candidates can delete older tweets (or indeed, entire campaign accounts),

and have done so more often for older campaigns compared to recent campaigns. Recent changes to the Twitter API have made it almost impossible to reliably find and track deleted tweets by candidates and elected officials (Willis 2023).

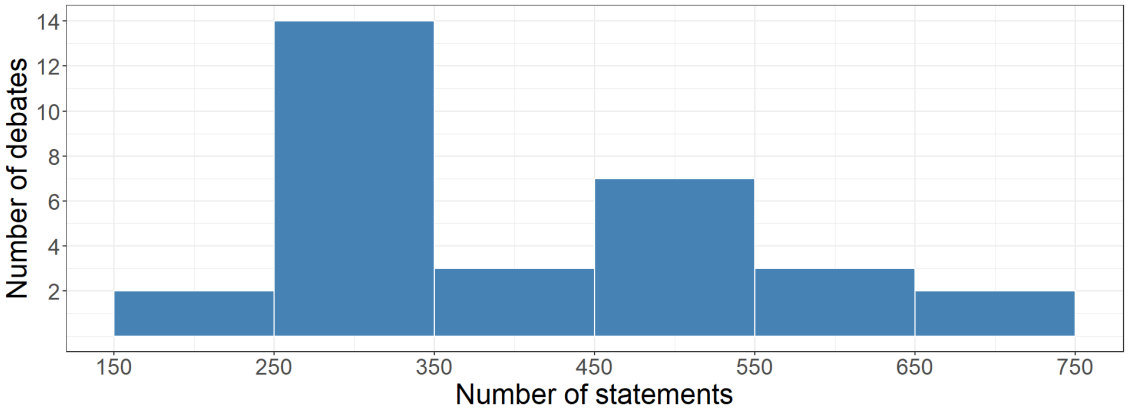
Debate statements

The debate search process outlined above yielded 31 debates, covering almost all (91%) of sampled elections. After separating individual candidate "speeches" into distinct sentences, I am left with 9,680 debate statements. The fact that every election had at least one debate was a function of the data collection process (see Table 1.1). I suspected that the debate transcripts would have better coverage of statements put out by front-runners than the Twitter sample given that lesser-known candidates sometimes do not appear in debates, especially if the candidate pool is large. While there are many more tweets than debate statements, the context of the debate statements means they have increased relevance about how candidates presented themselves to voters during the campaign. While candidates may occasionally tweet about topics unrelated to the campaign, statements delivered on a debate stage are more likely to be related to the campaign and thus more relevant for the present study of candidate messaging strategy.

The mean number of words in a debate statement was 25. I operationalize "statements" as full sentences or a sentence fragment for cases where a candidate began a sentence and was interrupted. Figure 1.7 displays the distribution of statements among the debates in my sample. The tight clustering of this distribution at the lower end reflects the fact that a majority of debates (19 out of 31) were broadcast on television. Moderators during televised debates exert control over the exchanges between candidates to fit the debate into a pre-established timeslot, usually sixty or ninety minutes. Moderator control thus limits the number of words that can possibly be

spoken in a televised debate and introduces many words (those spoken by the moderator) that must be deleted from the sample of candidate statements.

Figure 1.7: Distribution of debate statements in RCV elections



Note: Number of statements per debate calculated as number of sentences (or sentence fragments) said by all candidates during the debate, excluding moderator statements and one-word responses or utterances

Table 1.3 provides a broad view of the campaign rhetoric data and shows the number of statements associated with each election, separated by statement type. Most of the statements in my sample (67.7%) are tweets, which is unsurprising given that debate statements are subject to more restrictions (e.g., televised timeslot length, moderation interruption, speaking time limits). In the sample of 146 unique candidates, only 66 (45%) appeared in debates *and* used Twitter in the same election. One may wonder whether these 66 candidates were more "viable" than candidates who only used one type of messaging and are thus more worthy of study. However, that conclusion is incorrect for three reasons. First, the truly nonviable candidates were eliminated during the sample collection process. None of the 108 candidates that made no recorded statements came close to winning their elections, and these candidates often earned less than 1% of the total vote share. Second, some of the candidates ran before the use of Twitter was widespread; to exclude them from a sample of "viable candidates" would be misleading. For example, R.T. Rybak served four terms as the mayor of Minneapolis and won its first RCV

election by more than 60 percentage points, but did not create a Twitter account until 2010. Third, the meaning of "viability" is more ambiguous in a mixed sample of local and federal elections. In federal elections and high-profile primaries, it is often trivial to determine which candidates are "viable" based on their party affiliations, name recognition, or credentials. But as was mentioned previously, local elections have a much lower barrier to entry than federal elections.

Table 1.3: RCV campaign messages by type, year, and election

Year	Election	Debate Statements	Tweets	Total
2009	Minneapolis Mayoral	181	30	211
2010	Oakland Mayoral	0	780	780
2011	San Francisco Mayoral	456	2123	2579
2013	Minneapolis Mayoral	309	1339	1648
	St. Paul Mayoral	525	82	607
2014	Oakland Mayoral	746	495	1241
2015	Portland Mayoral	0	180	180
	San Francisco Mayoral	506	163	669
2016	Berkeley Mayoral	540	45	585
2017	Minneapolis Mayoral	587	1184	1771
	St. Paul Mayoral	701	458	1159
2018	Maine 2nd District General	284	488	772
	Maine 2nd District Democratic Primary	339	104	443
	Maine Senate	307	465	772
	Maine 1st District	246	397	643
	Oakland Mayoral	643	627	1270
	Santa Fe Mayoral	543	134	677
	San Francisco Mayoral	476	238	714
2019	Portland Mayoral	412	138	550
2020	Maine 1st District	276	247	523
	Berkeley Mayoral	286	651	937
	Maine 2nd District General	289	295	584
	Maine 2nd District GOP Primary	327	533	860
	Maine Senate	289	1826	2115
2021	Minneapolis Mayoral	303	1059	1362
	Santa Fe Mayoral	422	125	547
	St. Paul Mayoral	596	76	672
	NYC Mayoral Democratic Primary	471	6271	6742
	NYC Mayoral GOP Primary	294	272	566
2022	Alaska House Special	257	664	921
	Maine 1st District	291	267	558
	Maine 2nd District General	319	1637	1956
	Maine 2nd District GOP Primary	0	11	11
	Oakland Mayoral	386	2949	3335
	Total	12607	26353	38960

Note: Rows correspond to the total number of statements in each election. "Debate Statements" refers to individual sentences in televised debates. "Tweets" refers to individual tweets issued by candidates and includes retweets.

Additionally, most non-incumbent mayoral candidates in these elections have similar credentials, name recognition, and even party affiliation. To take one example, four of the six candidates for St. Paul mayor in 2017 were former city council members from the same political party (DFL). The small margins of victory in many municipal elections after multiple rounds of RCV tabulation is an indication that many voters consider a wide range of candidates to be "viable." Finally, candidates are always incentivized to portray themselves as a viable choice; no one wins votes by making it seem as if they will lose, even under RCV. Untangling projected viability from actual viability would require in-depth background research on each of the 146 candidates. Conducting that research would become even more difficult as the dataset expands and does not scale even to the present sample without a dedicated team of qualitative researchers. Thus, I set aside the question of candidate viability for the present study, though I acknowledge that future research on RCV elections may benefit from an extended study of candidate quality,

Conclusion

This dataset represents the largest existing effort to collect information related to candidate rhetoric in RCV elections in the United States. These data advance the scholarship on RCV, electoral systems, and campaign rhetoric in three key ways. First, I gather data on more RCV elections (n = 34) and candidates (n = 146) than previous studies, which often seek to "match" a small set of RCV and plurality elections. In doing so, past studies sacrifice the ability to generalize their findings across contexts and limit the relevance of their findings to the small set of elections that they survey.

Second, I collect information on the actual messages that candidates put out on Twitter and in debates, a method that is suited to my candidate-driven theory and allows for comparison of messaging across platforms. My multi-media method is different than previous work, which

has summarized the campaign "discourse" in the form of newspaper coverage *of* candidates or used surveys to gauge how voters react to candidate messaging (John and Douglas 2017).

Finally, to analyze the effects of individual electoral features, I collect more nuanced data on individual elections than past researchers, including the presence of ranking restrictions, historical competitiveness, the number of candidates, and the presence of an incumbent. These data form the basis of my multilevel sentiment analysis model that I employ in the following chapter to determine whether candidate and electoral features condition the moderating effect of RCV on campaign negativity.

Chapter 2: Testing the conditionality of the "moderating incentive" of RCV

In this chapter, I evaluate my hypotheses using a dataset composed of candidate statements on Twitter and in public debates. Using a set of three multilevel models and two methods of sentiment analysis, I find support for hypotheses related to gender, incumbent status, and incumbent presence. However, I fail to reject the null hypothesis for other variables tied to electoral administration. I extend the conclusions of previous research on plurality elections to RCV elections and show that female candidates and incumbents tend to engage in less negative campaigning than their male and challenger counterparts (respectively). My results also suggest that the moderating effect of RCV, which incentivizes candidates to engage in less negative campaigning, is robust across administrative and electoral contexts. This last finding is of particular concern given the decentralized nature of U.S. electoral administration.

In my introduction, I presented six directional hypotheses regarding the factors that condition the level of negativity in RCV elections. Based on the electoral systems and RCV literature, I predicted that female candidates and incumbents would engage in less negative campaigning than male candidates and challengers (H1, H2). I also predicted that RCV races with more candidates would have less negative campaigning than races with fewer candidates (H3), and that ranking restrictions would be associated with increased negative campaigning under certain conditions (H4). Finally, I predicted that open-seat and competitive RCV elections would exhibit less negative campaigning than uncompetitive elections and elections in which an incumbent was running, respectively (H5, H6).

My ability to evaluate these hypotheses is dependent on two items: access to a comprehensive set of campaign communications and a reliable method of measuring negativity. I addressed the former criterion in Chapter 1. In the following sections, I describe how I measure

campaign sentiment and outline why my approach is appropriate for my particular data and research questions.

Theoretical Overview: Campaign Sentiment

Tweets and debate statements

I operationalize the level of negativity in RCV campaigns by performing sentiment analysis on a combined sample of tweets and debate statements. This is an ideal approach that expands on past scholarship in several ways. First, scholars have examined the sentiment embedded in both tweets (Gross and Johnson 2016; Ceron and d'Adda 2016; Bekafigo and Pingley 2015) and debate statements (Marien et al. 2020; Hopmann et al. 2018) in isolation. Those methods have been empirically validated on recent campaigns within the United States and cross-nationally, and can be used to differentiate messaging and rhetorical styles with a high degree of accuracy (Karlsen and Enjolras 2016). That said, the use of only one medium may skew sentiment estimates and lead to a biased view of campaign civility due to incentives inherent to each medium. For example, some scholars suggest that platforms like Twitter may incentivize candidates to "go negative" more often than they would otherwise (Mueller and Saeltzer 2022). Additionally, while both Twitter posts and candidate debates are, by their nature, accessible to the public, citizens who engage in politics online may differ from individuals who engage offline, effectively constituting a different audience for candidates based on the medium in which their message is embedded (Preoţiuc-Pietro et al. 2017; Rainie et al. 2012).

By using campaign communications from multiple media, I also ensure that the dataset represents the full range of messages voters could have been exposed to in the campaign. Even at the local level, online campaigning has the potential to reach and influence more voters than appearances at public candidate forums (Keller and Königslöw 2018; Bode and Dalrymple

2016). The accessibility of online campaigning is of particular concern in RCV elections, which grant lower-performing candidates the ability to appeal to voters as a second, third, or lower-order choice (Donovan et al. 2016). The low barrier of entry to using Twitter for campaigning means this study captures messages from candidates not "viable" enough to make it onto the debate stage. However, these candidates may still affect the outcome of the election through the allocation of votes in additional rounds of tabulation.

Exclusion of advertisements

I analyze television advertisements from RCV elections separately in Chapter 3 with a modified case-study method, as advertisements differ from tweets and debates statements in several important ways. The additional work required to produce an advertisement—including filming, post-production, and purchasing airtime—creates distance between the message candidates want to convey and the final product. Television ads also represent a more targeted form of campaign messaging than tweets and debate statements because they appear on specific programs in specific regions at specific times. Scholars have shown that candidates can and do run ads designed to persuade specific subgroups of voters, rather than the entire electorate (Fowler et al. 2021a; Ridout et al. 2012). By contrast, tweets and debate statements are readily available to all voters. Finally, the high cost of television advertisements relative to other forms of messaging means that ads are differentially available to candidates in a non-random fashion. The problem of distribution is particularly acute for the present study given that most of the elections in my sample are local races, where candidates have fewer resources to finance television ads (Fowler et al. 2021b).

Another reason to analyze the ad data separately relates to data availability. Detailed ad data exist for only a handful of cases, almost all at the federal level, which makes it challenging

to draw general inferences about candidate strategy. Competitors in federal races have more resources than mayors of all but the largest U.S. cities—resources that can be used to create print, television, and online advertising. Candidates for the four federal offices in my sample ran a disproportionate number of available advertisements for my sampled elections, so much so that any generalizations are all but impossible. 41,063 of the 54,038 television ads in my sample (76%) were aired during six Maine elections for which data are available. While other RCV elections, including the 2021 NYC mayoral race, are of comparable size, prominence, and competitiveness to federal elections, ad data for those races will not be made public until early 2025. Future studies of RCV elections may benefit from an extended investigation of advertisements, especially low-cost methods like digital and online advertising.

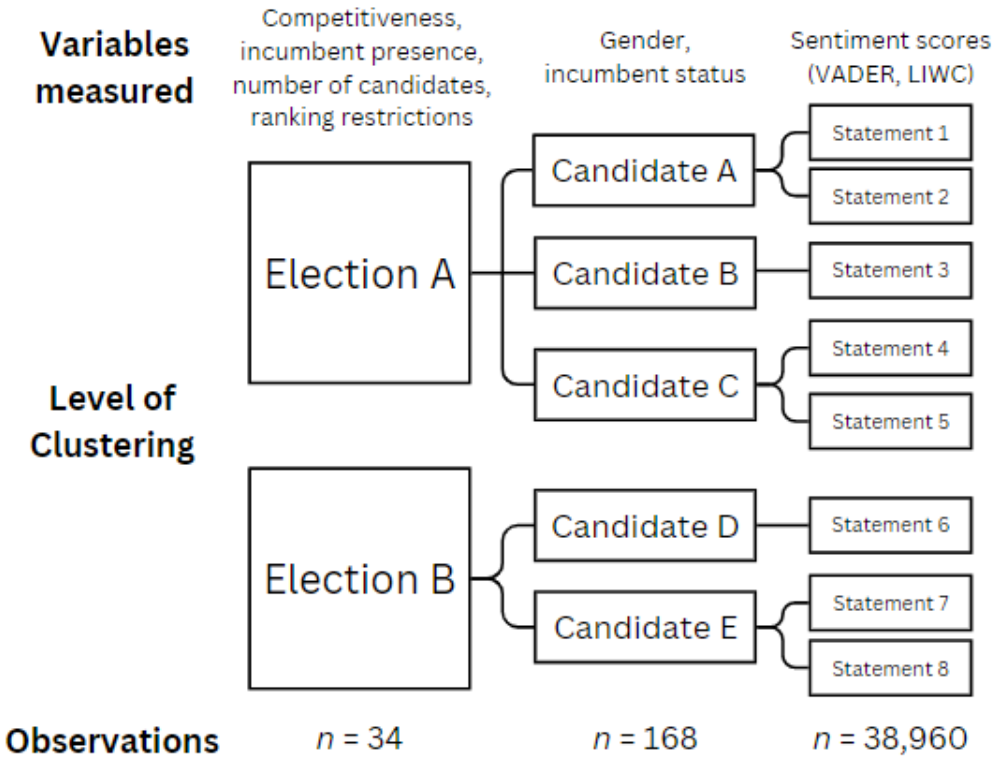
Methods

Model structure

I test my hypotheses using a multilevel regression model that regresses campaign negativity—as measured by sentiment analysis—on my key independent variables. A multilevel model is necessary given the nested structure of my data, which produces clusters at two levels. Each tweet or debate statement is associated with a specific candidate, who is in turn associated with a specific election. While the variation on the dependent variable (sentiment) occurs at the level of the *statement*, the variation in the independent variables occurs at the level of the *candidate* or *election*. This means the values of the independent variable(s) are clustered at the candidate and election level. To take one example, each statement made by Angus King, one of the two U.S. Senators from Maine, receives a sentiment score based on its content. But regardless of their content, all statements made by King have the same value for the independent variables *Incumbent Status* (i.e., 1) and *Gender* (i.e., 0).

Running the "typical" OLS regression model fails to account for clustering along candidate and electoral features. To deal with the nested and clustered nature of my data, I use the multilevel model depicted in Figure 2.1. Candidate statements (n = 38,960) are nested within candidates (n = 168), who are in turn nested within elections (n = 34). Model 1 (reported below) illustrates the structure of my data for all statements and elections in the sample.

Figure 2.1: Multilevel model structure



Note: Figure displays each level of clustering, the variables measured at each level, and the number of observations or clusters at each level.

There are many cases where RCV may not have influenced candidate strategy based on contextual factors that the dataset does not capture. For example, several races that nominally used RCV only contained two candidates. Under these circumstances, one candidate is guaranteed to win a majority of votes, and since no other candidates can be eliminated through instant runoff, the election essentially functions as a plurality election. In those circumstances,

candidates would feel no pressure to moderate their message based on the presence of RCV. Given that I am investigating the *conditionality* of the moderating effect of RCV on campaign negativity, I construct two alternative models that account for contextual factors by restricting the sample of candidates and elections. In effect, these models attempt to show evidence of conditionality where the moderating effect of RCV would be strongest.

Thus, in Model 2, I only include elections with three or more candidates and cases where ranking restrictions was "effective", i.e., when the total number of candidates exceeds the number voters were allowed to rank. Given that my ranking restrictions variable is dichotomous, I also include races without ranking restrictions to draw comparisons between the presence and absence of ranking restrictions. Model 2 includes 34,414 statements, 126 candidates, and 26 elections.

Model 3 is the most restrictive specification. In addition to the restrictions from Model 2, I only include elections that proceeded to multiple rounds of RCV tabulation. In other words, the only elections in Model 3 are elections where no candidate won an outright majority in the first round—the examples that advocates of RCV are most likely to mention. I expect the moderating effect of RCV on negative campaigning to be strongest in these cases, especially compared to less competitive elections. This portion of the analysis contains 21,983 statements, 83 candidates, and 13 elections.

Operationalizing independent variables

Independent variables in the present analyses fall into two categories based on the level at which they exhibit variance: election-level and candidate-level. For each race, I first record the *Number of Candidates* in the election. Given the wide range of my data along this variable (2 to 35), I also record the logged number of candidates. I then operationalize *Incumbent Presence* as

a dichotomous variable, taking a value of 1 when an incumbent was running and 0 in an open-seat race. As mentioned in Chapter 1, I also operationalize *Ranking Restrictions* dichotomously due to the low level of variance among ranking restrictions and the small number of sampled elections. Cases where voters were restricted to a specific number of candidates receive a value of 1, while elections with no limits receive a 0. I expect that ranking restrictions will act as a constraint on the moderating effect of RCV on candidate negativity. Finally, I operationalize *Competition* in the race as the mean of the margin of victory in the two previous election cycles:

$$Competition = \frac{M_{t1} + M_{t2}}{2}$$

Where *M* is equal to the margin of victory, calculated as follows:

$$M = Winner\ Vote\ Share\ (\%) - Second\ Place\ Vote\ Share\ (\%)$$

In all cases, the vote shares used to calculate competition are those of the *final round* of voting. RCV elections that progress to multiple rounds of vote tabulation often have smaller margins of victory than elections where the candidate won outright in the first round. However, this is a consequence of the design of RCV, and these closer margins reflect the outcome of the race.

At the candidate level, I operationalize both variables of interest dichotomously. Candidates receive a 1 on the *Incumbent Status* variable if they were in office during the election, and 0 otherwise. Given that the findings regarding gender and campaign sentiment focus on the differential effects of being a female candidate, I code female candidates as "1" on the *Gender* variable and code all other candidates as "0."

Dependent variables

Before conducting sentiment analysis, I engage in text-processing to ensure maximum performance of the respective sentiment models. I remove Twitter-specific symbols, acronyms, and nonwords from all tweets (e.g., mentions that include "@", hashtags, "RT"), as well as a set

of English-language "stopwords" that are both common and neutral (e.g., "I", "do", "the"). A full list of these deleted words and symbols can be found in the online supplementary material. 102 messages were removed during text processing because they contained only non-substantive elements. These include tweets only made up of emojis and retweets that only include mentioning (with "@") the handle of another Twitter user. Minor data losses of this type are a necessary cost of preprocessing. Additionally, to separate the data into meaningful tokens and to ensure that enough debate statements are present in the sample to draw meaningful conclusions, I separate the diarized debate transcripts at the sentence level. I perform the sentiment analysis using the "vader" and "quanteda" packages in R, which correspond to my two selected methods of sentiment analysis.

To measure campaign sentiment, I perform a dictionary-based sentiment analysis using the Valence Aware Dictionary for sEntiment Reasoning (VADER) and the 2015 Linguistic Inquiry and Word Count (Pennebaker et al. 2015; Hutto and Gilbert 2014). LIWC is widely regarded as the "gold standard" for dictionary-based sentiment analysis in the social sciences (Bryan and Ringsmuth 2016; Tumasjan et al. 2011). However, while LIWC can accurately gauge the underlying linguistic constructs in a text and displays predictive and concurrent validity (Bantum and Owen 2009; Kahn et al. 2007), the method has limitations for the present study. Specifically, research has shown LIWC performs better on longer text corpora and often struggles with shorter texts made up of fewer words (Tauschizk and Pennebaker 2010). Neither the Twitter statements, which are short by design, nor the debate statements, which have been segmented at the sentence level, are ideal for LIWC. To deal with these shortcomings, I primarily measure campaign sentiment using VADER, which uses a slightly different methodology than LIWC and has a high degree of performance on shorter social media data such

as Twitter posts (Elbagi and Yang; Ramteke et al. 2016). VADER was developed with applications for microblogging and social media posts, and it captures facets of social media discourse that LIWC cannot categorize, including acronyms, slang, and initialisms, which are key to understanding the sentiment of social media text (Davidov et al. 2010).

VADER analyzes text sentiment by rating individual words based on comparisons to a standardized "dictionary" of words, which was constructed using a variety of different types of text (Hutto and Gilbert 2014). VADER rates a word as "positive" or "negative," then provides a valence estimate between -4 and 4 for that indicates the strength of the sentiment of the word, with higher values indicating a more positive word (Hutto and Gilbert 2014). Once all words are rated, VADER produces three subscores—positive, negative, and neutral—for the entire text. These subscores represent the normalized sum of their respective word ratings, and have a range of -1 to 1, with higher values indicating that a text is more positive, negative, or neutral based on the particular subscore. The three subscores are then used to calculate a *compound sentiment score* which has a range of -1 to 1. On this final score, higher values indicate that the text is more positive, while lower values indicate the text is more negative. Completely neutral texts (i.e., texts that contain no positive or negative words) receive a score of 0. The compound score thus captures both the emotional valence of a text and the polarity of its sentiment, and it will be the basis for my analyses.

LIWC also uses a dictionary-based approach to measure campaign sentiment but has slightly different methodology than VADER. LIWC scans texts and assigns values to words that correspond to different categories, including part of speech identification (i.e., verbs, pronouns, prepositions) and a number of psychological constructs. In total, LIWC grades texts along 74 "linguistic dimensions" (Pennebaker et al. 2015). In this analysis, I record the values for the

following five dimensions: positive emotion, negative emotion, anger, anxiety, and sadness. Positive and negative emotion as measured by LIWC relate directly to the level of negativity in any statement and have served as the basis for prior work on this topic (Crabtree et al. 2020). "Anger," "anxiety," and "sadness" are labels that are applied to words that fall into the "negative emotion" category, and they provide additional nuance regarding *how* negativity manifests in a given text (Pennebaker et al. 2015).

LIWC evaluates texts along the dimensions listed above by matching individual words with words in its training dictionary. If a word that signifies "positive emotion" or "negative emotion" is present in the text, LIWC notes that correspondence. Importantly, the positive emotion and negative emotion lists are mutually exclusive; words appearing on one do not appear on the other. To calculate the overall "score" for a text along each dimension, LIWC sums the number of words that belong to each category and divides that sum by the total number of words in the document, generating a percentage score for each dimension (Pennebaker et al. 2015). I operationalize a Sentiment measure using the positive emotion and negative emotion percentages, calculated as follows:

$$Sentiment = \%\ Positive\ Words - \%\ Negative\ Words$$

I measure sentiment using a comparison of percentages based upon similar studies of campaign sentiment based on public statements (Crabtree et al. 2020). A higher Sentiment score indicates a more positive statement, while a lower sentiment score indicates a more negative statement. I also retain the measurements of positive words, negative words, and sub-categories of negative words (anxiety, anger, sadness) for potential exploratory analyses.

By using both VADER (numerical comparison) and LIWC (word categorization), I can be confident in my estimation of which words drive negativity in each statement (De Choudhury

et al. 2013). Unless otherwise noted, the analyses in this chapter employ the VADER measures. I include corresponding LIWC analyses in the online supplementary material. Table 2.1 provides examples of LIWC and VADER scoring for positive, neutral, and negative statements.

Table 2.1: Examples of campaign statement sentiment scoring

Candidate	Election	Statement	VADER sentiment score	Positive words (LIWC)	Negative words (LIWC)
Art Chang	2021 NYC Democratic primary	"Just for the record, @Taylorfor2021 is a great partner. She's **strong** but **humble**, **passionate** and **clear**, a **good** listener, but **strong** to her beliefs. It's a **pleasure** to campaign together—and compete! Rank @Art4MayorNYC 1 and @Taylorfor2021 2."	0.98 (extremely positive)	12.50%	0%
Nicholas Begich	2022 Alaska congressional race	"I went down to DC in March of last year to take a closer look at what was happening in the office and understand Congress a little bit more."	0 (neutral)	0%	0%
Aidan Hill	2020 Berkeley mayoral race	"I **hate** to bring this up on twitter but Wayne Hsiung is a **disgusting** human. I was sexually **assaulted** in his organization DxE (a cult) and hes writing articles saying I was an **abuser**. My body was **violated**. I was hospitalized. Wayne **disgusts** me."	-0.98 (extremely negative)	0%	11.54%

Note: Bolded words were selected by VADER and LIWC as containing sentiment. VADER sentiment score measured -1 to 1, where higher values indicate more positive statements. Statements made by Art Change and Aidan Hill sourced from Twitter without editing. The statement made by Nicholas Begich is sourced from a debate.

The statement made by Art Chang in Table 2.1 is an example of a strongly positive statement as scored by VADER and LIWC, in this case a positive evaluation of his opponent Joycelyn Taylor. VADER and LIWC identify similar lists of words (bolded) that express positive emotion (LIWC drops "passionate"). VADER then assigns sentiment to the individual words, as follows: "great" = 1.55, "strong" = 1.15, "humble" = 3.6, "clear" = 2.4, "good" = 2.85, "strong" = 3.45, "pleasure" = 4.0. The compound sentiment score, calculated by normalizing the sum of

each word listed above, is equal to 0.98, almost the highest possible score (1). The LIWC calculation is much simpler: the number of positive words (6) is divided by the total number of words (48) to produce a positivity score of 12.5%. Given that there are no negative words in the statement, the negative score is 0%. Simple arithmetic gives the Sentiment score of this statement (per LIWC) as 12.5 – 0 % = 12.5%.

The attack mounted by Aidan Hill against Wayne Hsiung, an electoral rival, is coded in a similar manner, save for the fact that the models identify it as almost entirely negative. Both VADER and the LIWC identify the same six words as being linked to negative sentiment. VADER assesses the words are as follows: “hate” = -1.35, “disgusting” = -3.6, “assaulted” = 3.6, “abuser” = -3.9, “violated” = -3.6, “disgusts” = -3.15. While not all of these words seem to directly express emotion, the training data VADER is based on indicates that words like “abuser” and “violate” are strongly linked to negative emotions. The intense degree of negativity present in these key words (almost all are near the maximum amount of negativity, -4) indicates just how negative the sentence is; the compound sentiment estimate given by VADER is -0.98. LIWC marks the same six words as negative and divides the number of negative words by the total number of words in the statement (52) to yield a “negative” words score of 11.54%. Given that LIWC identifies no positive words in the statement, the Sentiment score of this statement according to LIWC is -11.54%.

In contrast to these sharply emotional statements, the “neutral” statement offered by Nicholas Begich contains no words flagged by either dictionary tool as positive or negative. Thus, the Sentiment score calculated using LIWC data for this statement is 0, as is the VADER score. Additional confirmation with the VADER subscores (positive/negative/neutral) makes the case for emotional neutrality even clearer. While both the positive and negative subscores are

equal to 0, the normalized sum that informs the neutral subscore is equal to 1, the maximum possible value. “Neutral” statements are far more common in the sample than the emotionally charged examples given above.

Modeling campaign sentiment

After generating the sentiment scores for all statements (n = 38,960) in my sample, I regress sentiment score against the independent variables using a multilevel model with bootstrap clustered standard errors. The number of observations in my dataset is large at the statement level. However, running the regression model using standard errors at this level would be inappropriate given that the data is highly clustered; variation on the independent variables occurs at the candidate and election level.

Given that my data exhibits a high degree of clustering and that the sample size for the levels at which clustering occurs are much smaller than at the statement level, I use bootstrapping along with clustered standard errors to ensure sufficient statistical power. I employ clustered standard errors because the variance on the independent variable is not normally distributed (Abadie et al. 2023). In situations where the number of cases is small, as it is with both the candidate (n = 168) and election (n = 34) levels, scholars also recommend improving error estimates with bootstrapping (Cameron et al. 2008). Bootstrapping increases the number of observed clusters by resampling the data many times and generating new coefficient estimates and errors. Bootstrapping is particularly effective for nested models with few clusters at each level of the data (Huang 2018). Bootstrapped clustered standard errors have been successfully employed for similar analyses of campaign sentiment that used natural language processing and text analysis (Crabtree et al. 2020).

In addition to the variables described above, I also specify an interaction term between candidate gender and incumbent status, the two candidate-level variables in my analyses. Both gender and incumbency have been studied extensively in the negative campaigning literature (Auter and Fine 2016; Lau and Pomper 2004). Research has indicated that candidate gender *may* interact with other aspects of candidate identity (e.g., partisanship, race, ideology, personality), but no study of candidate rhetoric has directly examined the interaction between gender and incumbent status (Nai 2020).

Descriptive Statistics

I begin by presenting descriptive information about the specifications of my multilevel model and the independent variables used to test my hypotheses. Regarding the dependent variable, the overall levels of sentiment present in the dataset are similar across model specifications. The mean compound VADER score in Model 1 for all statements was 0.25, with a median of 0.30 and a standard deviation of 0.26. This is similar to Model 2 (mean = 0.25, median = 0.32, SD = 0.46) and Model 3 (mean = 0.26, median = 0.32, SD = 0.45). Additionally, in all three models, the VADER and LIWC sentiment measures are highly correlated at each level of clustering, as reported in Table 2.2.

Table 2.2: Correlation coefficients of LIWC and VADER sentiment measures

Cluster level	Model 1	Model 2	Model 3
Statement-level	0.65	0.64	0.65
Candidate-level	0.75	0.81	0.79
Election-level	0.71	0.75	0.68

NOTE: All correlations in Table 2.2 are significant at $p < .01$. VADER measure refers to the compound score generated by the VADER output. LIWC sentiment measure refers to the Sentiment measure derived from the LIWC "positive word" and "negative word" percentages as described above.

The similarity and strength of the correlations in Table 2.2 have two significant implications. First, the high degree of correlation at the statement level suggests that LIWC and VADER yield similar sentiment ratings for individual statements despite their different procedures and specifications. Second, the similarity between LIWC and VADER also is apparent at each level of clustering, which indicates that despite their different scales, the aggregated scores for candidates and elections reflect similar levels of sentiment under each method.

To demonstrate external validity, I adapt the methods of Crabtree et al. (2020) and compare the average sentiment of the short statements in my dataset (Tweets and debate statements) to other types of speech. Given that LIWC is based on a dictionary that explicitly draws on many types of writing, I calculate and compare the average amount of positive and negative emotion as scored by LIWC (Pennebaker et al. 2015).

Table 2.3: LIWC sentiment ratings by text type

	Combined	Sampled Tweets	Debates	Expressive Writing	LIWC Tweets	Natural Speech
Positive words	4.00	4.23	3.40	2.57	5.48	5.31
Negative words	0.98	0.97	1.02	2.12	2.14	1.19

Note: Values represent percentages of positive or negative words present in statements in my dataset or in LIWC training texts. Columns 1-3 contain my calculated values. Columns 4-6 are derived from other studies.

The first three columns in Table 2.3 display the percentage of positive and negative words as scored by LIWC in my dataset, breaking the total down by tweets and debates. I draw comparison between these data and three additional categories: “expressive writing,” which encompasses writing that explicitly has a strong emotional motivation, a sample of tweets surveyed by LIWC, and “natural speech,” which encompasses the emotional extremes present in interviews and in observed conversations (Pennebaker et al 2015).

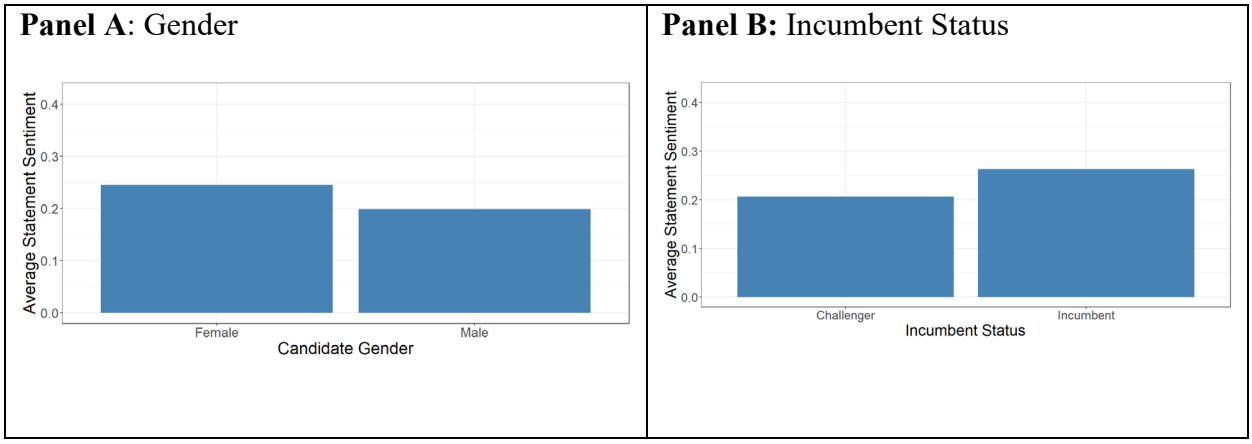
On average, the statements in my sample are more positive than they are negative. In addition, statements are substantially less negative than the other types of rhetoric that was part of training data for LIWC. Tweets are more positive (i.e., less negative) on average than debates. Additionally, these comparisons reinforce the idea that while debate statements and tweets may undergo review and development by campaign staff, they closely resemble natural speech with regard to the density of emotional expression.

Results

Bivariate analyses

I show bivariate relationships between the independent variables and average sentiment scores (measured by VADER), using all the observations at my disposal (i.e., Model 1). In most cases, the independent variables exert very small effects (< ¼ standard deviation) on campaign sentiment. The bivariate analyses depicted in Figure 2.2 seem to confirm my hypotheses regarding candidate-level variables.

Figure 2.2: Candidate-level variables and campaign sentiment

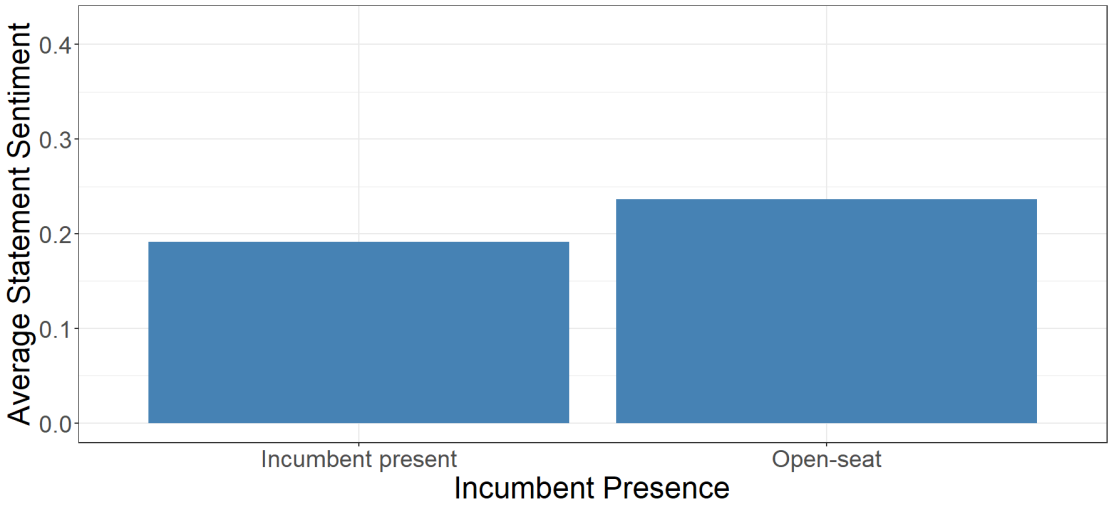


Note: Panels represent separate descriptive analyses of campaign sentiment and independent variables. Higher values indicate a statement is more positive and exhibits less negativity.

Female candidates in RCV elections are more positive (i.e., less negative) than male candidates, and incumbents engage in less negative messaging than challengers. However,

incumbent positivity is not sufficient to reduce the overall negativity of a race, as shown in Figure 2.3. The rhetoric in races with an incumbent is more negative on average than open-seat races. This may be because candidates in open-seat races all are less well-known to voters than incumbents, and thus can rely on weaker preferences on the part of voters to gain second- and third-choice support. Alternatively, the presence of an incumbent may stimulate negative attacks by challengers who believe that they can gain visibility by mentioning the more well-known incumbent in their statements and advertisements.

Figure 2.3: Incumbent presence and campaign sentiment



Note: VADER sentiment measured -1 to 1. Higher values indicate a statement is more positive and exhibits less negativity. "No restrictions" includes only elections that did not specify how many candidates voters could rank.

Interpreting the effect of ranking restrictions in Figure 2.4 is more complex, even for such a small change between the two groups. Races with ranking restrictions are slightly more negative than unrestricted races. This pattern suggests that where ranking restrictions apply, candidates are less constrained by the moderating incentive of RCV. Thus, candidates in races with restrictions will be able to engage in more negative messaging because they do not experience as strong of a moderating incentive as candidates in unrestricted races.

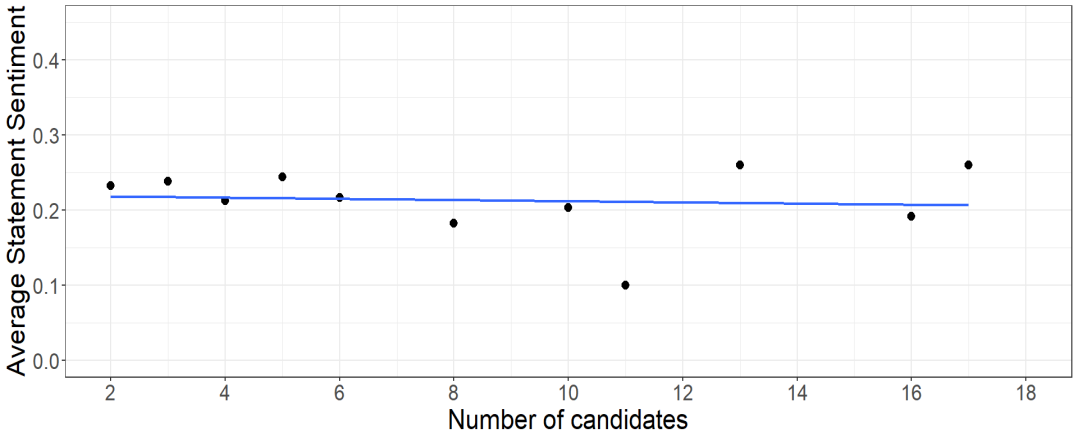
However, these bivariate analyses do not distinguish between races where restrictions were effective and races where they were not. By "effective," I mean races where the number of candidates exceeded the number specified by ranking restrictions, forcing voters to not rank some number of candidates. These races contrast with restricted races where the number of candidates was lower than the number specified by ranking restrictions, allowing voters to rank all candidates as if the race was unrestricted. I account for the "effectiveness" of ranking restrictions in my multilevel models, as indicated below.

Figure 2.4: Ranking restrictions and campaign sentiment

Note: VADER sentiment measured -1 to 1. Higher values indicate a statement is more positive and exhibits less negativity. "No restrictions" includes only elections that did not specify how many candidates voters could rank.

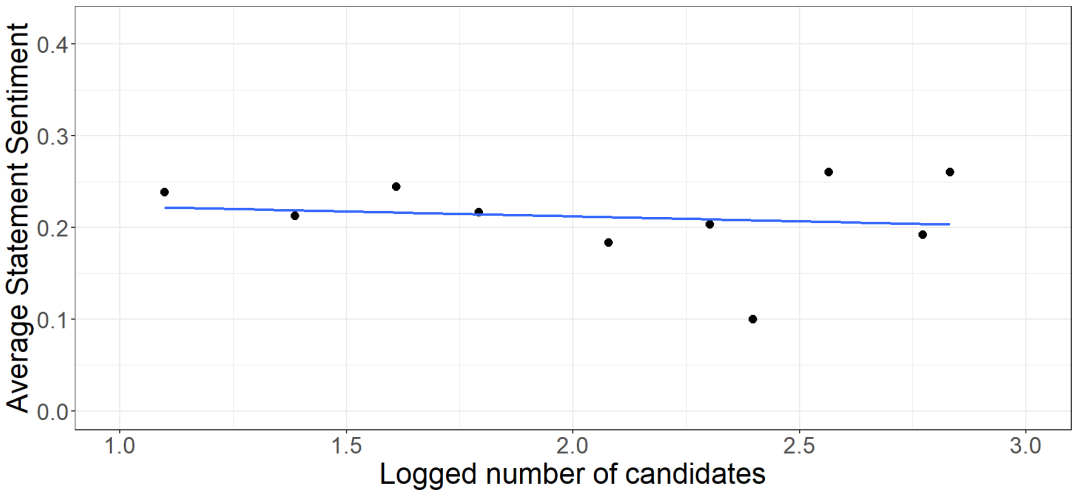
Figures 2.5—2.6 demonstrate that there is no clear pattern to the effect that number of candidates or logged number of candidates has on the average sentiment of campaign statements. It is worth noting the small number of elections in the sample and the discrete nature of the Number of Candidates variable. Any variation may be attributable to individual characteristics of the election rather than a general principle of elections with eleven candidates. Additionally, the distribution of candidates for elections in the sample is clustered at the lower end (i.e., most races have five candidates or fewer), further limiting the implications of this finding.

Figure 2.5: Number of candidates and campaign sentiment



Note: VADER sentiment measured -1 to 1. Higher values indicate a statement is more positive and exhibits less negativity. Trendline indicates line of best fit (residuals).

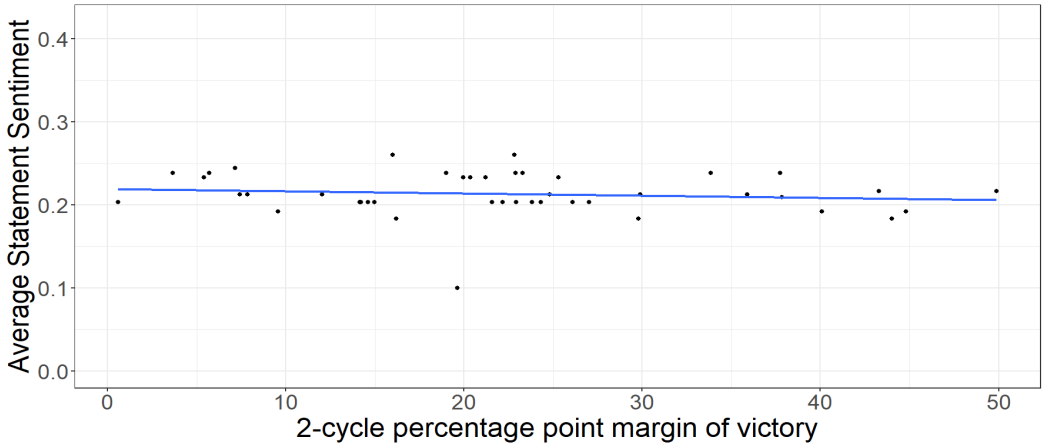
Figure 2.6: Logged number of candidates and campaign sentiment



Note: VADER sentiment measured -1 to 1. Higher values indicate a statement is more positive and exhibits less negativity. Trendline indicates line of best fit (residuals).

Figure 2.7 demonstrates a similar lack of a clear relationship between two-cycle percentage point margin of victory (i.e., historical competitiveness) and average campaign sentiment. This distribution highlights that regardless of competition, the RCV elections in the present sample were almost universally slightly positive, averaging to a 0.2 on the VADER sentiment measure.

Figure 2.7: Historical competitiveness and campaign sentiment

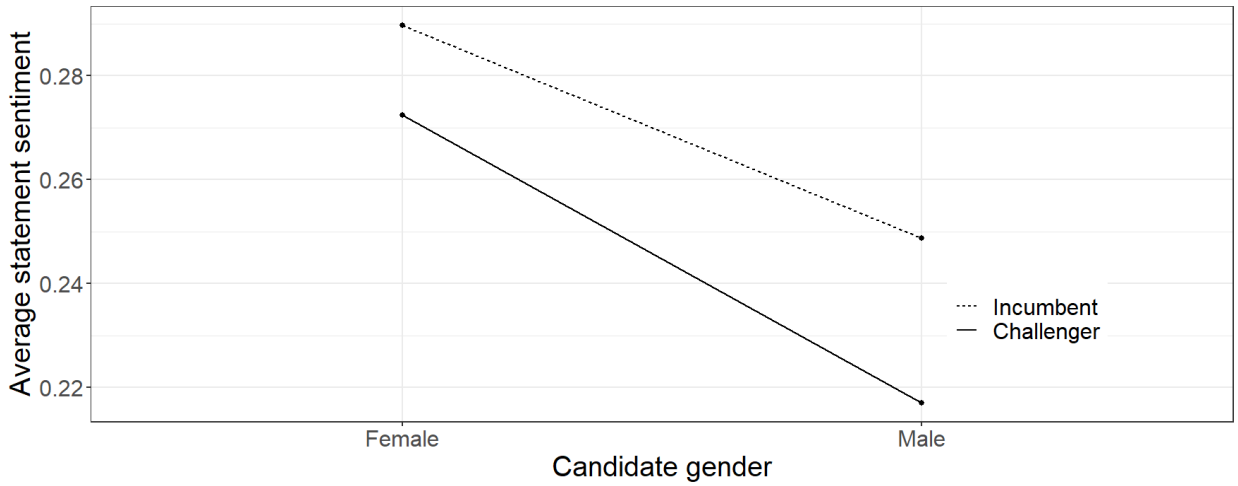


Note: VADER sentiment measured -1 to 1. Higher values indicate a statement is more positive and exhibits less negativity. Trendline indicates line of best fit (residuals).

Finally, Figure 2.8 demonstrates that I do not find evidence of an interaction effect between candidate gender and incumbent status, both of which vary at the candidate level. Female candidates are more positive (i.e., use less negative campaigning) than male candidates, and incumbents are more positive than their challengers. The nearly-parallel lines in the descriptive interaction plot indicate that the effects of gender and incumbent status are independent of one another. Additionally, the plot indicates that male incumbents are more negative than female challengers, another indication of the independent effect of each variable. Given that both effects operate in the same direction (i.e., pushing candidates away from negative campaigning), I expect that female incumbents will be the least negative type of candidate. However, female incumbents would not be *uniquely* less negative; their use of very little negative campaigning is a product of their gender and incumbent status acting separately, rather than an emergent property of the presence of both candidate-level traits. Similarly, I expect male challengers to be the most negative candidates, a product of the independent additive effects of their gender and status as a challenger (i.e., not an incumbent). All of these

effects are marginal, as indicated by the small differences between group means in Figure 2.8, but the independence of each variable as indicated by the interaction plot is robust nonetheless.

Figure 2.8: Gender and incumbency interaction



Note: Results of a descriptive comparison of the interaction between gender and incumbent status. Average VADER sentiment measured from -1 to 1.

Multilevel models of campaign sentiment

While the bivariate results above show small differences between conditions for all candidate-level and election-level features that I study, the results of the multilevel sentiment analysis model make it clear that only the candidate-level features have significant effects on campaign message sentiment. Table 2.4 summarizes the results of the multilevel models.

The effects of incumbency and gender are statistically significant, substantively small, and in the expected directions. I hypothesized that incumbents would be less likely to engage in negative campaigning, because of their increased visibility and ability to point to past policy accomplishments. In Models 1 and 2, incumbents have a 14 percentage point higher compound sentiment score than challengers, implying that they are slightly less negative (or more positive) than their challengers. Female candidates in RCV elections are slightly less negative (or more positive) than male candidates in RCV elections, on the order of 10-14 percentage points depending on the restrictiveness of the model. This significant effect supports my hypothesis

about the effect of gender on campaign rhetoric. The confidence intervals for both variables are relatively large across models, but the coefficient estimates remain statistically significant.

The effect of incumbent presence and incumbent status are not significant in Model 3, likely because of reduced variation on these variables within the most restrictive model. Among the 83 candidates in Model 3, only four (4.8%) were incumbents. Bootstrapped methods fail to generate appropriate confidence intervals for incumbent status and presence because resampling data with such little variation induces multicollinearity.

Table 2.4: Multilevel models of campaign sentiment

	Model 1	Model 2	Model 3
Candidate-level			
Incumbent Status	0.07*	0.07*	0.05
	(0.06)	(0.05)	(0.05)
Female	0.05*	0.07*	0.08*
	(0.05)	(0.05)	(0.07)
Cluster N	168	126	83
Election-level			
Number of Candidates	−0.00	−0.00	−0.00
	(0.01)	(0.01)	(0.02)
Ranking Restrictions	−0.01	0.00	0.00
	(0.05)	(0.08)	(0.06)
Incumbent Presence	−0.05*	−0.06*	−0.07
	(0.05)	(0.06)	(0.08)
Competition	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
Cluster N	34	26	13
Total N	38960	34414	21983

Note: Three-level multilevel regression model with bootstrapped clustered standard errors (β= 1000). * = $p < 0.05$ (two-sided). VADER sentiment measured -1 to 1, where higher values indicate more positive campaign rhetoric. Cluster N indicates the number of clusters at campaign or election level for each model, while total N indicates number of statements for each model.

The only election-level variable to achieve statistical significance is incumbent presence, which is associated with a 10-12 percentage point increase in average statement negativity

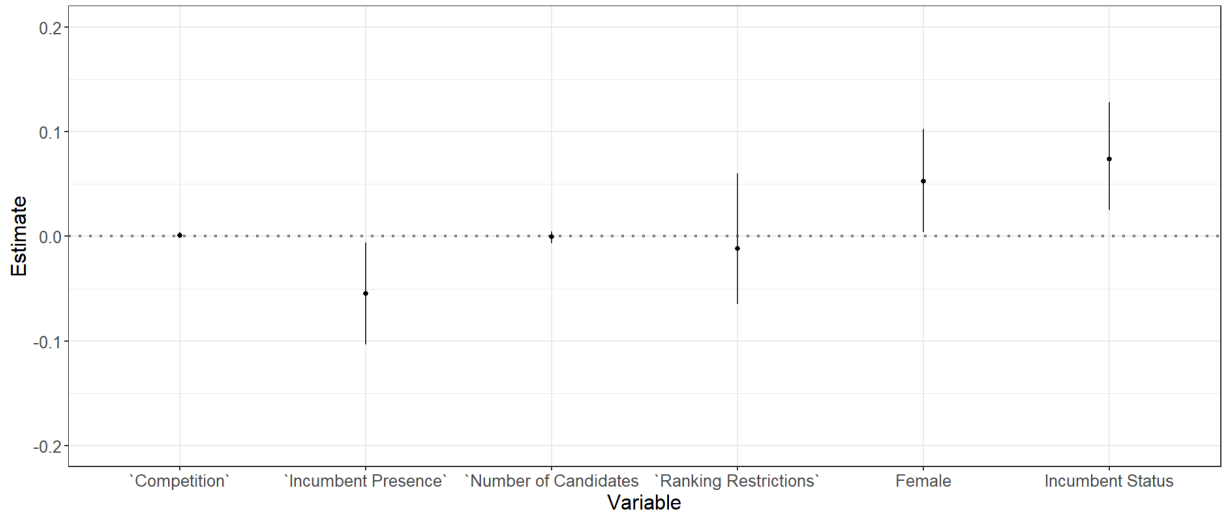
compared to open-seat races. The confidence intervals for incumbent presence are similarly large compared to the candidate-level variables, which is notable given that candidate-level and election-level variables were resampled equally (β = 1000 resamples). We would expect to find correlation between incumbent presence and incumbent status, given that at least one candidate in an election with an incumbent must be said incumbent. While both the candidate- and election-level version of the incumbency variable condition the level of negativity in RCV elections, the effects are in opposite directions. Incumbents may be more positive than their challengers, but elections with an incumbent are also more negative than open-seat races.

None of the other election-level variables reach statistical significance, and all estimates indicate that these effects would be very small even if they were significant. Further, the reported 95% confidence intervals indicate that if the effects of other election-level variables *were* significant, they would likely be substantively small. The restrictions I applied in Models 2 and 3 do not change the direction or strength of the effect of Incumbent Presence, though this pattern must be viewed with caution. The effects of ranking restrictions, competitiveness, and number of candidates are so marginal across the three models that even if they achieved significance, the negativity of campaigns across different conditions would be virtually indistinguishable.

Figures 2.6–2.8 are graphical representations of the results presented in Table 2.4. Across Models 1 and 2, the effects of gender, incumbent status, and incumbent presence are statistically significant ($p < 0.05$). The effects of candidate-level variables are positive in both models, while the effects of incumbent presence are negative in both models. All variables are of similar strength across models, though the confidence intervals for Model 2 are slightly larger, likely because of the reduced number of candidates and elections in the second model. In conjunction with the results in Table 2.4, these representations indicate that gender, incumbent status, and

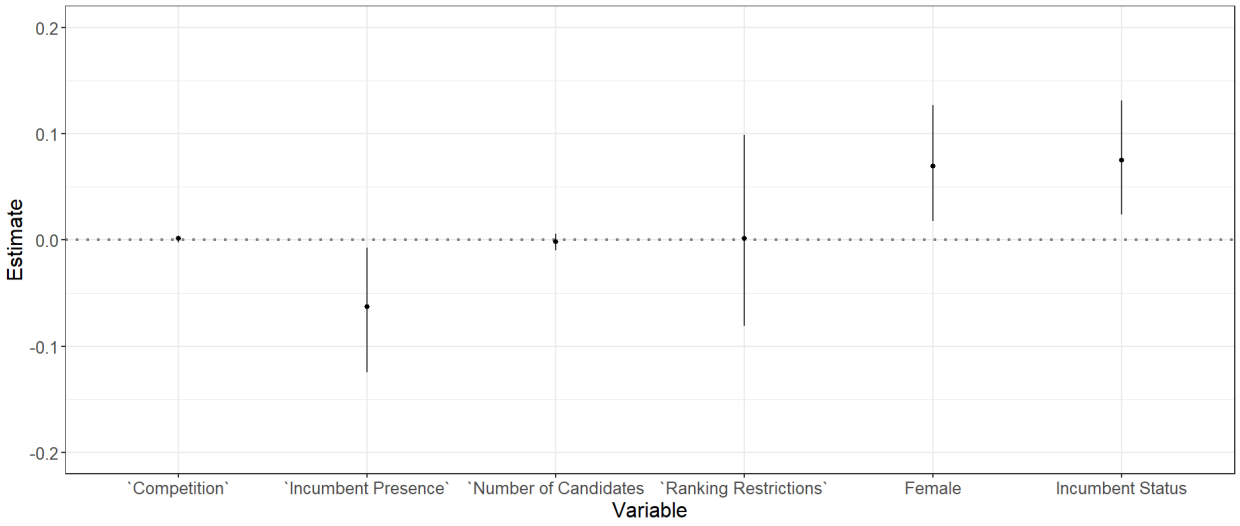
incumbent presence condition the amount of negative campaigning in RCV elections. Model 2 indicates that the above finding holds true regardless of whether ranking restrictions took effect or the number of candidates.

Figure 2.9: Effects of candidate and electoral features on campaign sentiment (Model 1)



Note: Lines represent 95% confidence intervals. Estimate represents coefficient estimate of change in average statement sentiment, measured -1 to 1.

Figure 2.10: Effects of candidate and electoral features on campaign sentiment (Model 2)

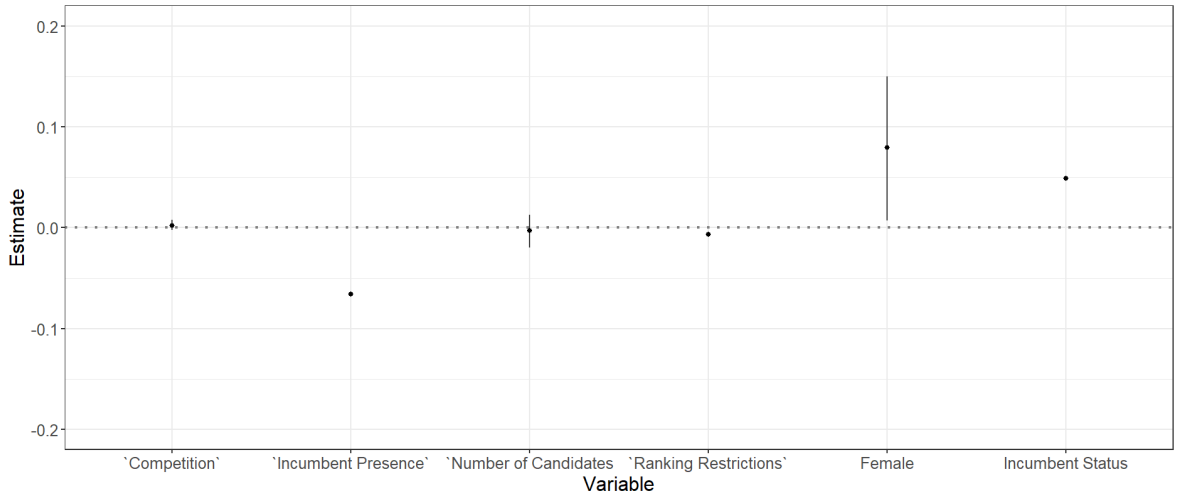


Note: CIs represent 95% confidence intervals. Estimate represents coefficient estimate of change in average statement sentiment, measured -1 to 1.

The restrictions applied in Model 3 include all restrictions from Model 2 as well as the specification that the elections must have proceeded to RCV tabulation (i.e., multiple rounds). This last restriction reduces the variance on several independent variables to the degree that even with bootstrapping, I cannot estimate reasonable confidence intervals. Figure 2.11 displays this finding. Most consequential RCV elections in the U.S. have not proceeded to multiple rounds of RCV tabulation; 19 out of 34 races in my sample (56%) were decided when a candidate won an outright majority in the first round. Multiple rounds of tabulation is the key feature that distinguishes RCV elections from plurality elections, and is central to my theoretical model. Given that quantitative methods are insufficient to compare single- and multiple-round elections, I return to the question of competitiveness and use of instant runoff for tabulation using qualitative methods in Chapter 3.

When I apply the models to subsamples that contain only debate statements or only Twitter posts, all independent variables fail to achieve statistical significance. Graphical representations of these findings are available in the online supplementary material.

Figure 2.11: Effects of candidate and electoral features on campaign sentiment (Model 3)



Note: CIs represent 95% confidence intervals. Missing Cis indicate a lack of sufficient variance. Estimate represents coefficient estimate of change in average statement sentiment, measured -1 to 1.

Discussion

The moderating effect of RCV on candidate rhetoric is conditional upon candidate gender (H1), incumbent status (H2), and the presence of an incumbent (H5). The direction of both candidate-level effects indicates that both factors enhance the "moderating effect" of RCV on negativity. The overall effect of RCV, compared to plurality/first-past-the post elections, is to incentivize candidates to reduce negative campaigning. The results of this study indicate that in addition to that baseline effect, incumbents and female candidates use even *less* negative campaigning than their male and challenger counterparts. This is an extension of the negative campaigning literature, which previously was concerned only with plurality elections. Furthermore, my analysis suggests that gender and incumbency condition the degree of rhetorical moderation independently. Thus, among RCV elections, which have a built-in incentive for candidates to reduce negative campaigning, we would expect female incumbents to be the most likely to moderate and male challengers to be the least likely to moderate their rhetoric.

Given that I examine only RCV elections, these data cannot be used to determine if campaigns in a plurality system will exhibit rhetorical moderation after implementing RCV. However, the candidate-level conditionality that I find indicates that studies of RCV moderation may find more complex outcomes. Changes in the use of negative campaigning may be due to individual differences at the candidate level rather than changes at the electoral level (i.e., from a plurality to RCV election). Additionally, the similarity of my findings across media (i.e., in debates *and* on Twitter) indicates that moderation is not confined to a single mode of communication. Instead, the decision to moderate extends across platforms and impacts all elements of the public persona associated with a candidate.

The effect of incumbent presence at the election level raises complex considerations, particularly when paired with the effect of incumbent status at the candidate-level. In line with previous research, my findings indicate that incumbents engage in less negative campaigning than challengers, but that negativity increases overall in races with an incumbent. These results support H5, which asserted that open-seat races would exhibit less negativity than races with an incumbent. The effects of incumbent presence and status cannot be compared directly given that the variation on both variables is present at different levels of clustering. However, these findings make intuitive sense given that an RCV election with three or more candidates will always have one incumbent and multiple challengers. Even if incumbents moderate their rhetoric and engage in less negative campaigning, the additive effect of multiple challengers going negative will push the overall tone of campaign rhetoric towards negativity.

The fact that the use of negative campaigning is conditional both on incumbent presence and incumbent status is of particular concern given that voters are sensitive to negative attacks by specific candidates *and* general levels of negativity in an election. Incumbents engage in less negative campaigning, and they may be able to avoid voter backlash by remaining positive. Additionally, given that incumbents are associated with the status quo, attacks by more negative challengers are likely to target incumbents and their policies. Thus, challengers may experience voter backlash even if they do not engage in egregiously negative or personality-based attacks. By this reasoning, the decision by a challenger to go negative is a particularly bad strategy in an RCV race, which already extends an incentive to moderate. Embracing negativity under non-optimal conditions may explain why incumbents succeed so often in RCV races. Challengers only won five of the 21 elections with an incumbent, a success rate of just under 24%.

In contrast to the significant effects of candidate-level variables, the effects of election-level factors are largely null, implying that these factors do not condition the rhetorical moderation under RCV. Nevertheless, my findings remain important for the future of electoral administration and RCV implementation. This study represents the most comprehensive overview of RCV elections in the United States to date, and the analysis incorporates campaign messages from nearly every major contemporary RCV election at the federal and local levels. The lack of conditionality based on election-level qualities displayed in these data is the strongest indication to date that the moderating effect of RCV on candidate messaging, sentiment, and civility is robust across administrative contexts.

The null effects of all variables in separate samples do not necessarily imply that the effects of RCV on campaign rhetoric are robust across different forms of media. Factors the condition the overall level of negative campaigning (e.g., gender, incumbent status, incumbent presence) fail to condition it in separated samples of tweets and debate statements. However, this may be due to lack of statistical power, especially for the smaller debate subsample. Further research using a mix of sources is needed to determine whether use of negative campaigning in RCV races by candidates is targeted or otherwise conditional on the media used.

The implication that the moderating effect of RCV is robust across contexts is of particular interest for implementation in other states and municipalities. Research shows that the difference in negativity between RCV and plurality elections is significant but substantively small (Kropf 2021; Donovan et al. 2016). Additional work has shown that voters are sensitive to reduced levels of negativity in RCV races (John and Douglas 2017). If RCV is robust across candidates and electoral contexts, this small effect on negativity will persist regardless of where RCV is implemented. Further, candidates who adjust their rhetorical strategy following RCV

implementation will reap the potential rewards from voters regardless of the context of their election. Of course, any modifications to their messaging will be at least partially dependent on their gender and incumbent status. Further still, differential implementation of RCV due to decentralized electoral administration (e.g., ranking restrictions in one city but not another) seems unlikely to affect candidate civility enough to affect the baseline effect of RCV.

My results should be interpreted with some caution given the construction of my sample. The fact remains that while the data used in this study represent the broadest examination of RCV to date, the number of cases at both the candidate and election levels in the multilevel models are relatively small. There have not been enough RCV elections in the United States to draw meaningful conclusions about electoral context, campaign messaging and candidate strategy, with standard levels of certainty or statistical precision. Additionally, I excluded many RCV elections from the final sample, as detailed in Chapter 1, and the absence of these data may not be independent of candidate strategy. For example, I dropped all local races below the mayoral level, but mayoral campaigns might have unique strategic pressures that are absent in higher-level races or in races for different offices at the municipal level.

While bootstrapping the clustered standard errors mitigates the problems of a small sample size, the process of resampling increases the amount of measurement error in the dependent variable, increasing the likelihood of a classic Type II error (Cameron et al. 2008). In other words, my null results may be a false negative, and election-level factors may indeed exert an effect on campaign sentiment. Additional exploratory analyses using only clustered standard errors indicate that this is likely the case—the resulting effects are larger and more significant than the results from the bootstrapped models. But drawing conclusions from so few clusters is

far more damaging to the robustness of the models than the problems introduced by bootstrapping.

Setting aside concerns about sample size and measurement error, a closer examination of the results indicates that the effects of the election-level variables other than Incumbent Presence would be small even if they had been significant. There is no guarantee that voters would respond to such marginal changes. While each variable could exert cumulative moderating effects in conjunction with others (e.g., competitive races without ranking restrictions vs. uncompetitive races with restrictions), my data do not permit an exploration of cumulative impacts.

While I can only show for certain that the effect of RCV is conditional on candidate-level variables, my results suggest that the conditional effect of electoral features is likely to be small and unlikely to negate the baseline moderation effect. The 95% confidence contains zero for almost all estimates (13 out of 18) across the models, which makes it unclear if the various independent variables exerted any effect on the level of negativity for each candidate or election. But even in the most extreme cases, the confidence intervals are narrow. This suggests that even in cases where I cannot be certain an effect exists, if a variable *does* influence campaign negativity, that effect is likely to be substantively small (Rainey 2014).

The similarity of the candidate-level effect sizes across all three models suggests that the impact of the moderating incentive of RCV is not dependent on whether candidates believe that RCV tabulation will occur, which is itself a function of competitiveness. The effects of gender and incumbent status were unchanged even when only examining elections with more than two candidates and where the counting proceeded to multiple rounds. Additionally, the insignificant results for historical competitiveness affirm the idea that the mere *presence* of RCV may exert a

moderating effect on candidate rhetoric across contexts, independent of whether RCV comes into play.

Conclusion

The results in this chapter lend support to three of my hypotheses: incumbents engage in less negative campaigning than challengers, female candidates engage in less negative campaigning than male candidates, and races with an incumbent tend to be more negative on average than open-seat races. However, the downstream effects of studying a small pool of elections and candidates make it unclear if the conditionality of RCV-related rhetorical moderation on electoral features is absent or if it is masked by measurement error.

The quantitative methods employed in this chapter have some limitations. The small number of ranked-choice elections in the United States makes drawing conclusions about these elections difficult, even with a comprehensive set of campaign communications. My data capture only a portion of full campaign context, as the tweets I include (the majority of the statements in the dataset) were made during the closing two months of the campaign seasons. Furthermore, my results may be vulnerable to omitted variable bias, as I examine a small number of individual and administrative features in complex, 21st-century elections. However, my selected variables all have precedents elsewhere in the negative campaigning or electoral systems literature. Finally, reconstructing candidate intention and strategy is a difficult task, particularly with a set of tools that were trained on a wide variety of text corpora.

Chapter 3 will attempt to address the shortcomings of this chapter using a qualitative case-study approach to study a small set of elections: federal contests in Maine after 2018. In doing so, I hope to capture candidate reactions to RCV implementation, incorporating the variables from Chapter 1 as well as campaign finance data and television advertisements. This

analysis will examine interaction between campaign strategy, message moderation, and electoral success in a general context, providing insight into the possible consequences should additional jurisdictions across the United States implement RCV.

Chapter 3: The Maine effects: case studies of campaign rhetoric

The previous chapter demonstrated that the level of negative campaigning employed by candidates in RCV elections is conditioned by candidate-level variation on gender and incumbent status. In those analyses, I employed a comprehensive dataset of statements drawn from a large sample of RCV elections. However, my tests only accounted for a limited set of independent variables at both the candidate and election levels. Additionally, while the data I drew upon were extensive, spanning 34 elections and more than 38,000 statements, the robustness of my conclusions remains limited by sample size. There simply may not have been enough US elections that use ranked-choice voting to make more nuanced inferences using aggregated measures of message negativity.

In this chapter, I adopt a qualitative approach and conduct a series of in-depth case studies of RCV elections. I leverage the availability of additional data for federal elections and explore the contextual factors that may contribute to strategic decisions that involve when and how to use negative campaigning. I compare three federal elections in Maine—the 2018 & 2020 Senate races and the 2018 race for the Second Congressional district—that serve as illustrative examples of how RCV influences candidate strategy. I find that reducing negative campaigning is only a worthwhile strategy under certain circumstances. Furthermore, my findings suggest that the moderating incentive of RCV may only impact electoral outcomes in districts that are already competitive.

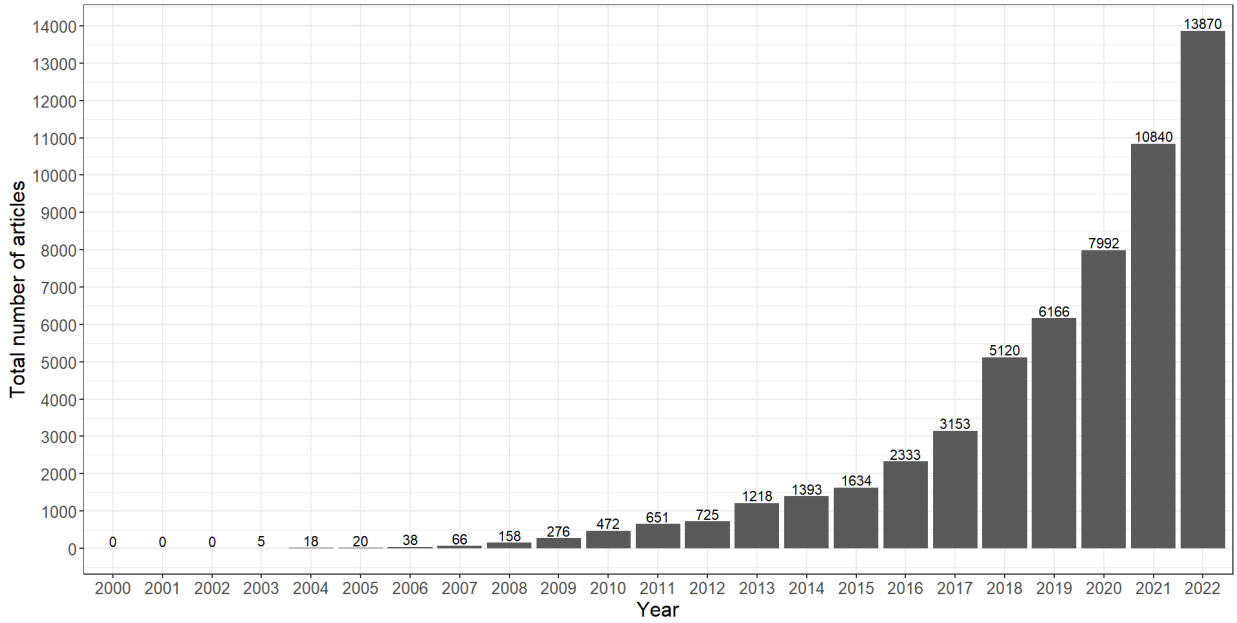
Data

Elections

By restricting the cases studied to a sample of federal elections, I concentrate on analyzing the effects of RCV on candidate strategy in contests that are widely publicized,

normatively important, and the most relevant to future discussions of RCV implementation. Politicians elected to municipal offices by RCV typically only have jurisdiction over local policy matters, but politicians elected at the federal level influence policy decisions that affect the entire nation. Additionally, U.S. states are more likely to adopt RCV or allow its passage by popular referendum if other states are able to implement it successfully, but are less likely to be swayed by municipal adoption in other states (Santucci 2018). Finally, the use of RCV to elect federal legislators attracts more media attention than RCV implementation at the state level. To illustrate this last point, Figure 3.1 displays the results of a search of the Nexis Uni database of U.S.-based newspapers and publications for the use of the term "ranked-choice voting."

Figure 3.1: News articles that mention RCV by year



Note: Results obtained by searching the Nexis Uni newspaper database for the term "ranked-choice voting" in U.S. publications, excluding duplicates.

Media attention to RCV grew meaningfully only after a 2016 referendum guaranteed that Maine would implement RCV by referendum in 2018, with the number of discussions about RCV reaching new heights after Alaska used RCV in the 2022 midterms. Federal elections are also worthy of study because of their precedence in the negative campaigning literature and the

wealth of additional data that are available compared to municipal elections, including polling, television advertisements, and campaign finance filings.

While federal elections are worthy of study, it would be a mistake to simply use the variables employed in previous chapters to assess the impact of RCV and electoral features on the usage of negative campaigning in these races. Table 3.1 details the federal general elections in my dataset and highlights what little variation exists on key independent variables.

Table 3.1: Characteristics of federal RCV elections

Election	# of candidates	Victor	Victor party	Margin of victory (percentage points)	# of rounds
2018 Maine Senate	3	Angus King	Independent	19.1	1
2018 Maine House 1	3	Chellie Pingree	Democratic	26.3	1
2018 Maine House 2	4	Jared Golden	Democratic	1.2	2
2020 Maine Senate	4	Susan Collins	Republican	8.6	1
2020 Maine House 1	2	Chellie Pingree	Democratic	24.4	1
2020 Maine House 2	2	Jared Golden	Democratic	6.0	1
2022 Maine House 1	2	Chellie Pingree	Democratic	25.7	1
2022 Maine House 2	3	Jared Golden	Democratic	6.1	1
2022 Alaska at-large House district	4	Mary Peltola	Democratic	10.0	3

Note: Margin of victory calculated from Ballotpedia and verified by electoral returns, using the final margin of victory in cases with multiple rounds of tabulation.

The number of candidates varies from 2 to 4. As discussed in Chapter 2, elections with only two candidates do not function as an RCV election. An incumbent was present in all races except the Alaskan at-large House District—and incumbents won without exception in all but

one case. No races had ranking restrictions, and while there was some variation on competitiveness, only two races advanced beyond the first round of tabulation. Studying the effect of RCV on candidate strategy in federal races clearly requires a different approach guided by new sources of data on my independent *and* dependent variables of interest.

Advertisements

In addition to the sentiment measures used in Chapter 2, I draw on the sentiment embedded in televised political advertisements, an approach that adds depth to the Twitter and debate data. While assembling the data used in previous chapters required extensive manual collection and quality-testing, the ad data in this chapter were sourced directly from the Wesleyan Media Project (WMP). The WMP uses a systematic approach with multiple human coders to perform in-depth content analysis on the raw data, which consists of video files of campaign advertisements. The raw data come from the Campaign Media Advertising Group (CMAG), which draws from all U.S. media markets to create a comprehensive dataset of campaign advertisements (Fowler et al. 2016, 9). These data include more variables than the previous aggregated measures of campaign sentiment computed by using VADER and LIWC.

From a functional perspective, advertisements are different than other types of campaign messaging. There is more "distance" between the candidate and the final product. Creating a political ad involves input from many actors depending on the size and status of the campaign. Further, the process of deploying an ad across even a single media market requires navigating interactions between media buyers and broadcast television representatives (Fowler et al. 2016, 88). Federal campaigns in the 21st century increasingly rely on media and political consultants to execute various steps of the ad production process, which includes background research, ad creation, focus testing, and final production (Johnson 2007). Larger campaigns may employ

many individuals to research opponents or members of their district, fine-tune elements of the ad (including music, color, and scripting), gauge constituent responses to the ad, and launch it in concert with the press, which may be considered a campaign message in and of itself (Fowler et al. 2016, 81). The amount of effort, coordination, and fine-tuning required to produce an ad is substantively different than tweets or debate remarks, even if those communications are also subject to staff oversight.

Competition reviewed: polling and campaign finance

By restricting my focus to a smaller set of federal RCV elections, I use additional data to better differentiate between competitive and uncompetitive elections. As discussed in previous chapters, the historical competitiveness measure was sufficient for the sentiment models despite its weaknesses. Percentage point gradations in past margins of victory are not the focus of this chapter. In any case, I am not limited to using the past margin of victory as an indicator of competitiveness. Instead, I can approximate the "perceived competitiveness" of each race by examining campaign finance filings that the FEC requires of all federal candidates and pre-election polling data, which both candidates and the public can readily access. Campaign finance toplines provide an indication of the total resources candidates have at their disposal. Further, candidate campaigning level of spending relative to their opponents can serve as a proxy for how much effort the candidate was putting into the race (Cox 2020). Additionally, more detailed campaign finance data contains information about individual disbursements, including the timing and size of media or advertising purchases.

Polling data is a more precise indicator than campaign finance data given that polls directly measure the preferences of voters. However, polls are conducted less frequently than the regular FEC filing deadlines. Opinion data also can be difficult to access, and survey-level

variation in question wording and sampling method may affect the perceived margins between individual candidates even before accounting for margin of error (Kennedy et al. 2018). A race may be considered competitive with a margin of between five and eight percentage points depending on the polling agency—a wide gap in an era of increasingly close margins at the federal level (Gonzales 2023; Li and Leaverton 2022; Fowler et al. 2016, 45). With these difficulties in mind, I assess competitiveness on a case-by-case basis, acknowledging that a race may seem more or less competitive at various points during the campaign, and may swing between both extremes even in the final two months of the race.

Qualitative case studies

The present study requires a qualitative approach owing to its small sample size and limited variation on relevant independent variables. The number of federal elections that have used ranked-choice voting is too small to draw meaningful conclusions using quantitative methods even with resampling or bootstrapping errors. As shown in Table 3.1, the federal elections in my sample do not exhibit enough variation on the independent variables of interest to draw robust conclusions. The problem of insufficient variance persists even when adding primary elections, some of which proceeded to two rounds of tabulation.

Importantly, however, there is variation across my cases on candidate-level factors and competitiveness. Using additional sources of data, I can make a more holistic judgement about the competitiveness of each federal race, approximating the information available to candidates that guided their strategy. Restricting the sample to a smaller group of candidates also allows me to form a more coherent profile of each candidate, rather than viewing them as the product of dichotomized factors. Understanding individual candidates is especially important given that I

am studying Maine, a state known for political moderates and successful independent candidates at the state and federal levels.

Given the data at my disposal, I analyze three federal RCV elections using a qualitative case study approach. The wealth of additional information about these elections permits me to investigate how RCV interacts with contextual factors beyond the variables from Chapter 2. I select three elections that illustrate the potential effects RCV may have on candidate strategy at the federal level. In Case 1, I discuss the re-election of Angus King to the U.S. Senate in 2018, a scenario in which RCV did not have a significant impact on candidate messaging strategy. In Case 2, I discuss the much more competitive Senate race in 2020, which ended with the re-election of Susan Collins. I argue that while RCV *may* have played a role in candidate strategy, it likely did not influence the outcome. Finally, in Case 3, I discuss the 2018 election in the Second Congressional District. In this case, I argue that the narrow victory won by Jared Golden in the second round of RCV tabulation was due in part to his choice to moderate his campaign rhetoric, a strategy that his opponent rejected outright.

Case 1: RCV and the lone Independent

The first case I examine is one in which the moderating incentive generated by RCV did not impact the outcome and was not taken up by any of the three candidates, including the victor, Angus King. In the 2018 Senate race, King successfully ran for re-election as an Independent against Eric Brakey, a Republican state senator, and Zak Ringelstein, a former schoolteacher and novice Democratic candidate. While Ringelstein, who finished a distant third, publicly incorporated RCV and its moderating incentive into his strategy, King enjoyed a massive incumbency advantage and did not have to moderate in order to win. Ultimately, I conclude that RCV did not affect candidate strategy in this race because its moderating incentive was swamped

by other contextual factors, including campaign spending and the incumbency advantage. Candidates acted according to other incentives because they did not perceive that the gains from moderating their rhetoric and appealing to other voters as a second- or third-round choice would be worth changing their messaging and not playing to their strengths and the situation.

King had no additional incentive to moderate his rhetoric in the first RCV Senate election ever conducted for two main reasons: he could be certain he would win a first-round majority and he had a history of using almost no negative campaigning. Before running for Senate in 2012, King had served two terms as the first Independent governor of Maine. After a narrow victory in 1994, where he beat his Democratic opponent by less than two percentage points, King went on to win a second term by a much wider margin, beating his 1998 challenger by forty percentage points while carrying every county in the state. Though he remained out of the political spotlight between 2002 and 2012, his popularity allowed him to defeat two state representatives and again win an outright majority to become a U.S. Senator. By the time his 2018 re-election campaign was underway, King was the most popular politician in Maine and heavily favored to win. A Suffolk University poll from August 2018 found that 63% of likely voters in Maine viewed King favorably, with only 27% holding an unfavorable view, and 52% indicated they would rank him as their first choice.

King enjoyed great popularity during the 2018 campaign thanks to his reputation as a moderate in Congress, which appealed to the Maine electorate and influenced his messaging strategy. In that same Suffolk University poll, likely voters in Maine split their partisan identification into roughly equal parts—34% said they were registered Democrats, 29% said they were registered Republicans, and 33% identified as unenrolled or Independent (Suffolk University 2018). Though King caucuses with Democrats, he had taken (and continues to take) a

middle ground as a policymaker in order to appeal to as many citizens as possible in his state. King has DW-NOMINATE score typical of a centrist, particularly when compared to other Democrats (Voteview 2018; Poole 2005). King has a record of being an outspoken moderate, as indicated by other metrics like bill sponsorship (GovTrack 2018).

King used his status as a moderate and incumbent to fuel a positive campaign in Maine, as reflected in advertisement and sentiment data. The WMP tracked 1002 television ads sponsored by King and his campaign, *all* of which were positive. These ads highlighted his past accomplishments, character, and future goals. Across Twitter and the debate stage, King had an average compound VADER score of 0.193, higher (i.e., less negative) than average for the 2018 Senate election (0.155), but slightly lower than the average for statements across the dataset of all major RCV elections in the U.S. (0.242). The difference between the average sentiment in the 2018 Senate race and across the sample of all RCV elections may be due to the actions of the challengers. The two opponents that King faced pushed the tone of the race in a more negative direction, usually by attacking him. By contrast, King did not even use his opponents by name during the debates or on Twitter, and when he brought up other parties, it was to stress his ability to compromise, as with the following quote from one of the statewide debates:

> there are lots of issues where I've been *able to work with both sides* starting with the student loan crisis in 2013 when the rates were just about to double a small group of us *Republicans Democrats and myself put together a compromise* to solve that problem (emphasis mine).

In sum, the electoral context for the 2018 race seemed ideal for King to engage in rhetorical moderation after RCV was implemented. However, he already was engaging in a campaign that did not use negative campaigning, and even if he did, his baseline popularity and appeal to Maine voters would still likely have guaranteed his victory.

The main source of opposition to King was two-term Republican state senator Eric Brakey, who did not attempt to moderate his rhetoric throughout the uphill battle of the campaign. Instead, Brakey leaned hard into his role as a challenger and engaged in repeated negative campaigning, courting a right-leaning base of support by attacking King from a conservative perspective. While all eighty of the TV ads that Brakey ran were positive, he routinely attacked King on both personal and policy matters, both on Twitter and in debates. The mismatch between the two front-runners in terms of number of ads was apparent in the campaign finance figures. Brakey spent just over $1.2 million during the campaign, in contrast to King, who spent $5.9 million, nearly five times as much. Brakey had a mean compound VADER score of 0.187, slightly less than that of King (0.193), and he attacked King by name ninety-four times in the two months leading up to Election Day. In addition to criticisms of past policy decisions made by King, including aid to Syria and voting for the Affordable Care Act ("Obamacare," in his words), Brakey called King a "fake Independent…bought and paid for by the energy lobby." Brakey went on to win 35.4% of the vote, finishing 19.1 percentage points behind King.

While the top-two vote-getters did not appear to be influenced by the moderating incentive of RCV, Zak Ringelstein seems to have embraced it, which is worth considering in further detail despite his poor poll performance. A public-school teacher, Ringelstein was a political novice that mounted a challenge against King from the left. He was the only major-party Senate candidate in 2018 to be endorsed by the Democratic Socialists of America (Stack 2018). Given that Ringelstein faced steep odds from the start, he acted as past literature would predict and went negative early and often. While Ringelstein only ran thirty-two ads, all of them were "contrast ads" that presented positive information about himself alongside negative "contrasting" information about another candidate. Additionally, he had the most negative mean VADER score

of the three candidates, at 0.09. Like other lesser-known candidates across the sample, the debate was a chance for Ringelstein to get more attention than he would otherwise, and he used his status as a challenger to launch negative attacks when given time to speak. The average VADER score of statements Ringelstein made on Twitter 0.11, but on the debate stage, it was 0.02, far lower than his two opponents.

However, Ringelstein differs from other lesser-known candidates that used negative campaigning at all times. Instead, he selectively moderated his rhetoric based on the candidate he was attacking, reserving direct attacks for Eric Brakey and using only qualified policy criticisms against Angus King. Ringelstein repeatedly called on Brakey to drop out of the race for promoting white nationalist rhetoric, centering his criticism on the hostility Brakey harbored towards refugees. During one of their debates, Ringelstein actually *defended* King in response to a negative attack by Brakey, taking the opportunity to fire off an attack of his own:

> we need to respect our Senator here. I mean, he's a good guy and he served our state well. Senator King, I thank you for your service, but I do think it is time for the future. And I do think it is really disgusting Eric that you've decided to go the way of the Trump agenda with violent rhetoric…

When he did criticize King, Ringelstein often added positive messaging to ostensibly negative campaigning, as in the tweet below, which was posted after King agreed to reject donations from fossil fuel companies:

> I pressured @SenAngusKing for months on this and now he's giving back fossil fuel money he took. This is the kind of bold leadership I will demonstrate in the US Senate every single day.

Finally, though he used negative campaigning against both Brakey and King, Ringelstein would often conclude his appeals (some of which were negative!) by calling on voters to use RCV in a specific way, namely, to rank him first and King second. The following is an illustrative example from a debate between the three candidates:

> …we need to push a vision that is not corporate, that is American, that is human. And the people of Maine deserve better. I'm asking for you to rank me one, Angus second, and I'm asking for you to do that because we can vote our hopes. *There is no spoiler effect.*

Here, Ringelstein alludes to the "spoiler effect," which refers to the principle that voting for third-party candidates is a waste of a vote and that third-party candidates can harm other candidates that they are aligned with. Ringelstein explicitly acknowledges the fact that no spoiler effect exists under RCV (since votes for losing candidates are redistributed in later rounds) while urging a positive vision of the future.

Ringelstein ultimately lost, earning just over 10% of the vote. However, his messaging style indicates that in cases where the moderating incentive of RCV is not overwhelmed by other factors like incumbent status, the way in which RCV interacts with negative campaigning may be more complex than initially thought. Candidates may choose to moderate their rhetoric selectively, adopting different tactics based on their visibility in different settings. Additionally, while candidates may use similar or elevated levels of negative campaigning, they may target their attacks selectively based on poll position and ideological proximity to other candidates. This last point may be significant for races where several co-partisans compete with one another, as is the case for many municipal RCV elections.

Case 2: Four-party split and the 2020 Senate race

The hints of complexity that were present in Case 1 are much stronger in Case 2, the 2020 Senate election in which Susan Collins, the incumbent, defeated Maine House Speaker Sara Gideon and two independent challengers to win re-election. The election did not proceed beyond the first round of tabulation, and the final margin of victory was 8.6 percentage points, outside the conventional upper bound of what is considered competitive. However, though the effect of RCV on the outcome is unclear, the presence of additional candidates and the *perceived*

competitiveness of the race created a situation in which some candidates embraced the moderating incentive of RCV.

Though Collins won re-election, the race was much closer than her previous 2012 victory, which she won by 36.2 percentage points. Collins was known as a moderate Republican largely because of her voting record, and she was popular in Maine in the years leading up to the election (Bolton 2020). An August 2018 poll from Suffolk University found that 49% of likely voter rated her as "favorable," with only 31% viewing her as "unfavorable." However, by the middle of the 2020 campaign, those values had nearly reversed. A June-August 2020 poll conducted by Quinnipiac University found that only 42% of registered voters viewed her favorably, while 49% viewed her unfavorably, and disapproved of her job as Senator by a nine-point margin. These abysmal numbers made her the most unpopular member of the Senate in her home state at the time (Bolton 2020). Indeed, many polls showed Gideon and Collins within five percent of each other—a September poll conducted by Quinnipiac University even showed Gideon ahead by eight percentage points.

Exactly *why* Collins seemed to be performing so poorly before the election but beat the odds and won is outside the scope of this thesis. That said, her weak pre-election polling likely reflected her high-profile support of an unpopular presidential administration and the perceptions that Gideon was a strong contender. Collins voted to acquit Donald Trump at his first impeachment trial, and she voted to confirm Brett Kavanaugh to the Supreme Court. These two actions undermined her past (and, by all accounts, desired) image as a pro-choice moderate who voted according to her stance issues, not mere party loyalty (Cauterucci 2020). Collins also faced a formidable opponent in 2020. Gideon was a high-profile candidate at the state level, having served four terms as a state representative and two terms as Speaker of the Maine House of

Representatives. Gideon also far outspent Collins in what was (at the time) the seventh most expensive Senate campaign of all time (Cilizza 2020). According to FEC filings, by the end of the race, Gideon had spent $62.9 million, which was more than twice what Collins spent ($29.6 million).

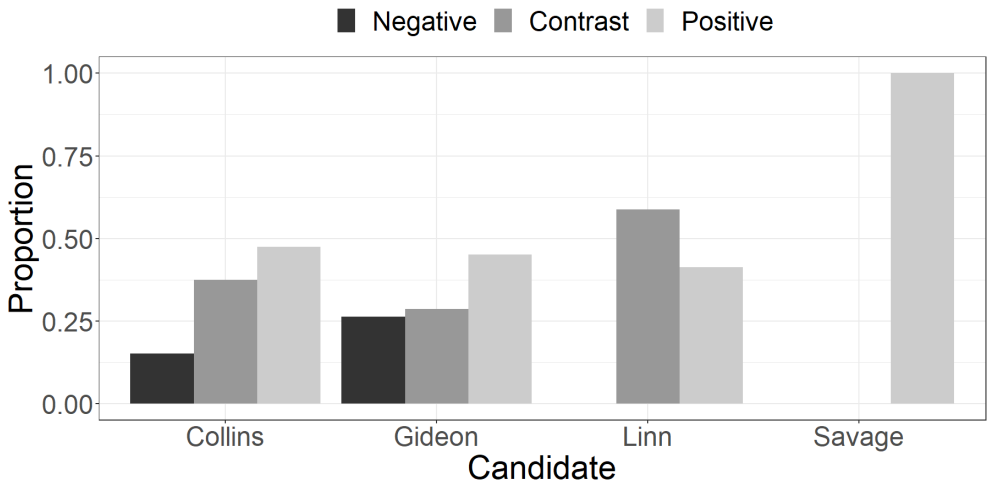
According to my theoretical model, it should come as no surprise that Collins, perceiving that the race might be close (and thus might go to multiple rounds), used almost no negative campaigning throughout the election. Judging by mean statement VADER score, the 2020 Maine Senate race was more positive than the average RCV race. The mean score for all statements in this election of 0.289, slightly higher than the to the overall average across all RCV elections (0.242). The fact that this race was more positive than average is unexpected given that Collins was an incumbent and there were numerous challengers, all of whom were expected to use negative campaigning more than her. The relative absence of negativity in this race serves as an indication that the strength and quality of the incumbent plays a role in how their presence affects the level of negativity. Collins had a mean VADER score of 0.404 across Twitter and debate statements, much more positive than her fellow candidates. However, Collins did not completely refrain from attacking Gideon, and often criticized her for wanting to "defund the police," "hike taxes…on heating oil," and for having "set a record in the amount of dark money that is flooded the airwaves in this state."

Despite being a challenger, Gideon likely also saw the race as potentially competitive, and may have sought to soften some of her rhetoric in a similarly complex way as Collins. Gideon had a mean compound VADER score of 0.304—lower than Collins, but still above average for the 2020 Senate election (0.289) and the full sample (0.242). However, she attacked Collins 224 times on Twitter and in debates, often using the same policy-focused framings that

Collins used to avoid alienating voters (e.g., criticizing how Collins fundraised, how she flip-flopped on issues, and how Collins supported Justice Kavanaugh).

Together, Collins and Gideon ran 35,200 television ads during the campaign, totaling more than half (51%) of all the ads I obtained from the Wesleyan Media Project for the elections in my sample. Unlike Case 1, the increased competitiveness of the 2020 race seems to have induced more negativity in campaign messages broadcast as ad spots, as shown in Figure 3.2.

Figure 3.2: Types of ads in the 2020 ME Senate race



Note: Bars represent the proportion of each type of ad used by each candidate. "Contrast" ads include negative information about opponents and positive information about the candidate, and are classified separately from negative ads by the WMP.

Both Collins and Gideon used about the same proportion of positive ads (Collins 47%, Gideon 45%). However, Gideon used a higher proportion of direct attacks, while Collins used a higher proportion of contrast ads. In conjunction with the rosy poll numbers Gideon enjoyed and her spending advantage, Gideon viewed the race as more in her favor (i.e., less competitive). Thus, she may have experienced a weaker incentive to moderate her rhetoric compared to Collins, the comparative underdog and incumbent who engaged in less negative campaigning and was more positive overall.

Analyzing the decision calculus of candidates in this race is complicated by the fact that there were two additional Independent challengers: Lisa Savage, who ran to the left of Gideon, and Max Linn, who ran far to the right of Collins. Figure 3.2 shows that Linn chose to engage in more negative campaigning than Savage; 59% of his 211 ads were contrast ads, while all 11 of the ads that Savage were positive. Linn regularly sparred with both his fellow candidates and even the moderator while on the debate stage, though his mean compound VADER score (0.209) was slightly higher than Savage (0.191). Both Independents offered critiques of their opponents, but Savage criticized Collins and Gideon on a myriad of policy issues, including campaign spending, environmental protections, and support for universal healthcare, while Linn offered both personal and policy criticisms (albeit much more infrequently).

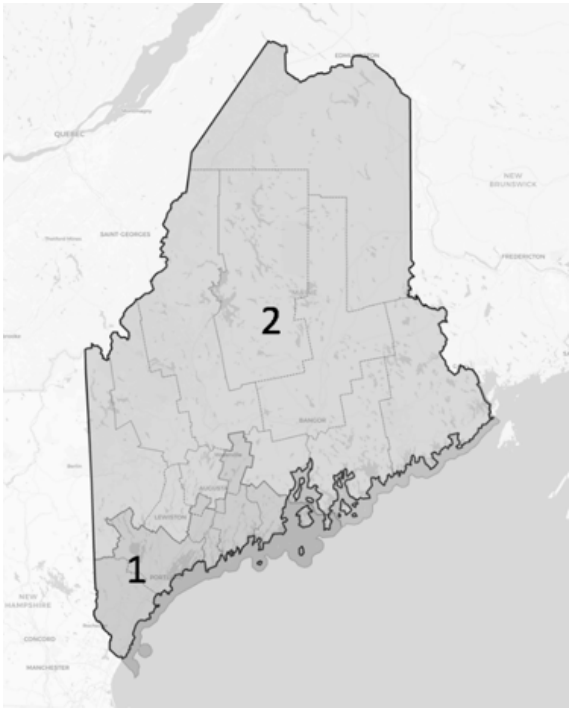
The decision on the part of both Independent candidates to "go negative" at a higher-than-expected rate in a competitive election may be because perceptions of competitiveness depend on candidate poll position. While the race might be competitive between two other candidates, third-party candidates may intuit that they will receive a reduced share of the vote even under RCV, and thus engage in more negative campaigning if they feel they cannot win in additional rounds. Thus, the principle that RCV reduces negative campaigning in competitive races may be more nuanced than a simple measurement of historical competitiveness can capture. While I chose not to study "viability" in quantitative terms, modeling candidate strategy may require an examination of candidate *perceptions* of their own "viability." I cannot show that Linn and Savage expected that winning the election even after additional rounds of tabulation, but if they had done so, they would have been right. While the moderating incentive may have come into play in the 2020 Maine Senate election, Susan Collins won in the first round with a slim majority—just 51% of the vote, according to the Maine Secretary of State.

Case 3: Golden vs. Poliquin & Baber v. Dunlap: House District 2 in 2018

In Case 1 and Case 2, the impact of the moderating incentive of RCV on candidate strategy was shown to be largely irrelevant for some candidates and more nuanced than my original expectations for others. Case 3 concerns perhaps the most controversial case of RCV in the United States—a situation where the incentive to moderate may have determined the outcome in the 2018 race for the Second Congressional District.

While all Senate seats are statewide races, the campaigns for U.S. House in Maine take place inside its two congressional districts, the boundaries of which are indicated by Figure 3.3.

Figure 3.3: Map of Maine congressional districts



Note: Regions represent districts as of September 2021.

District 1 encompasses much of the Portland and Augusta metropolitan area, the two major population centers in the state; the Democratic incumbent, Chellie Pingree, has handily won there each time she has been elected under plurality and RCV. While both House districts are of roughly equal population, District 2 covers a much larger geographic area, largely because of its

demographics. District 2 has both the second-highest concentration of non-Hispanic White residents (94%) and the second-most rural residents (72%) in the country. As Table 3.1 demonstrated, elections in the Second Congressional District have been more closely contested in recent years, owing to the conservative bent of its rural electorate. The closest of these elections was the 2018 general election—the first U.S. congressional race to proceed to multiple rounds of RCV tabulation, and thus far the only race to have caused an "upset" between rounds.

The incumbent and prospective front-runner was Bruce Poliquin, who had narrowly defeated progressive Democrat Emily Cain in 2014 and 2016, continuing a tradition of Republican control of the Second Congressional District. Poliquin held the view that RCV would not influence the outcome, and that it might not be legal (Collins 2018). Even after the 2018 primary, in which Poliquin ran unopposed, RCV was caught up in a set of state court decisions regarding usage in state-level elections (Maine State Legislature 2022). As the 2018 election moved into its final months, the future of RCV in Maine seemed uncertain.

Still, if Poliquin was paying attention to widely-publicized local and national polls, he would have viewed the election as competitive. Two polls conducted by the New York Times in the fall of 2018 showed his September lead of 47%–42%(± 4.8) shrinking. By October, Poliquin was in a dead heat of 41%–41%(± 4.8). Competitive conditions may have spurred him to spend more than he had previously, with his expenses totaling $4.3 million over the course of the campaign.

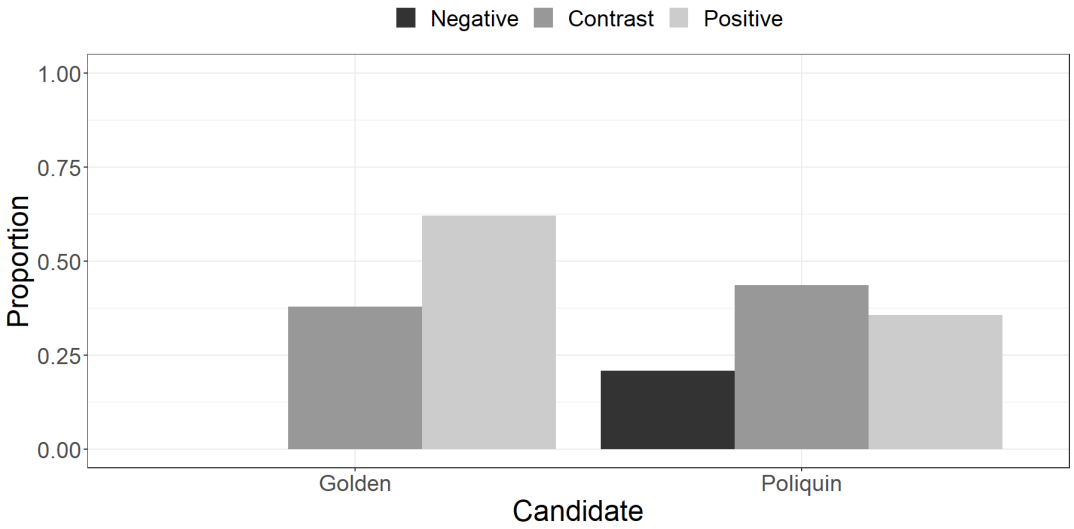
I would expect Poliquin to be motivated by his incumbent status and the state of the race as highly competitive to moderate his rhetoric under RCV, but he did just the opposite. Instead, Poliquin was a relatively negative candidate, employing tactics that had succeeded against state representative Emily Cain, his previous progressive challenger. The mean VADER score for

statements in the 2018 Second Congressional District race was 0.241, almost equal to the average across all major U.S. RCV elections (0.242). Compared to Cases 1 and 2, Poliquin enjoyed only a modest incumbency advantage, which may have mitigated the effects of his incumbency on his negativity. However, compared to his opponent, he did not choose to moderate. Poliquin had an average VADER score was 0.218, a sharp contrast to his opponent, Democrat Jared Golden, whose average statement scored above average (0.288) for this race and across all races in the sample.

Golden, the former Assistant Majority Leader in the Maine state legislature, had made his way through a contested primary (which proceeded to two rounds of RCV tabulation) by focusing on the economy and stressing his stance as a moderate on salient issues like gun control (Nemitz 2018). Golden knew he would have to fight an uphill battle, and outspent Poliquin by roughly $1.8 million. Throughout the four-primary and the four-way general election, Golden restrained his use of negative campaigning. Like Sara Gideon, Golden could rely on his past as a state legislator, pointing to past accomplishments and responding to attacks with policy criticisms, a method of messaging usually only available to incumbents. Poliquin often attacked Golden directly and personally, characterizing him as a "young radical" who "believes in the socialist agenda." In response, Golden would often stress his support for moderate reforms like Medicare expansion and compromise gun control, including expanded background checks. When Golden criticized Poliquin, it was along policy lines and unrelated to personal differences. Golden often brought up the votes Poliquin had made in Congress, including a vote to repeal the Affordable Care Act and a vote to pass the 2017 Tax Cuts and Jobs Act. Golden attacked this latter decision and claimed that the tax bill had benefited the wealthy, and mentioned both subjects in debates and on Twitter.

Both major candidates made extensive use of television advertisements, thought neither of the Independent candidates did. Though Independents Will Hoar (slightly right-leaning) and Tiffany Bond (slightly left-leaning) did appear in debates, their presence was relatively marginal, especially compared to the Independent candidates discussed previously. Figure 3.4 displays the distribution of ads used by Poliquin and Golden.

Figure 3.4: Types of ads in the 2018 ME Second Congressional District race



Note: Bars represent the proportion of each type of ad used by each candidate. "Contrast" ads include negative information about opponents and positive information about the candidate and are classified separately from negative ads by the WMP.

While Golden ran 1,520 more ads than Poliquin, he did not employ any direct attack ads and ran a lower proportion of contrast ads. Though Poliquin was the incumbent, he chose to promote his record less often than Golden, who could only rely on his accomplishments at the state level. The disparity in types of advertisement is another piece of evidence that Golden and Poliquin took sharply different messaging strategies. While Poliquin attempted to portray his opponent as a radical leftist, Golden responded with moderate policies and moderate policy criticisms, stressing his qualifications according to the moderating incentive of RCV.

While it is impossible to know for certain if Golden planned to win by appealing to second-choice votes, but the electoral returns indicate that second-choice support from Independents won him the election. Poliquin had been leading in the first round by approximately 2,000 votes but lost by just over 3,000 votes (1.2 percentage points) once the second-choice votes for Bond and Hoar were allotted to him and Golden. Poliquin went on to demand a recount and made a national media splash by filing an injunction in federal district court to have the ranked-choice results overturned, arguing he was the actual winner (Shepherd 2018). Whether Poliquin had planned to engage in legal action all along is uncertain. What is known is that before the month was out, Judge Lance Walker struck down the claim that RCV was unconstitutional in *Baber v. Dunlap* (Carman and Glushefski 2019).

In considering the outcome of this election for the purposes of this study, the broader national context should not be overlooked. The 2018 midterms were a landmark win for the Democrats, who rode a wave of suburban dissatisfaction with the Trump administration to victory (Jacobson 2019). Poliquin had wholeheartedly embraced the policies of the Trump administration, often defending the president and his economic record. While the Second Congressional district is predominantly rural, it also included a sizable proportion of suburban voters who overwhelmingly broke for Democrats in 2018. This race remains significant as a milestone for electoral reform and RCV implementation, but a different perspective could legitimately argue that it was just another of the forty-one House seats that Democrats won in the 2018 midterms.

Conclusion

The most significant development in the recent wave of RCV implementation has been the use of RCV at the federal level, beginning in Maine in 2018 and more recently in Alaska in 2022. Politicians elected using RCV are often more representative of their voters than those

elected by plurality (John et al. 2018; McGing 2013), and the use of RCV in higher-profile races has attracted attention from scholars and the media in recent years. In this chapter, I analyzed a subsample of RCV elections, focusing on three races at the federal level in Maine: the Senate elections in 2018 and 2020 and the 2018 election in the Second Congressional District.

The findings of my case studies suggest that the "moderating incentive" of RCV can play out in different ways based on the interaction between contextual factors. In Case 1, I explored how leading candidates in uncompetitive races may not experience as great an incentive to moderate. In such races, these candidates engage in less negative campaigning than those who trail in the polls. In Case 2, I found further evidence that moderation may be dependent on poll positioning, and described conditions by which the effect of RCV on candidates may have been overwhelmed by competing concerns, including incumbency and the makeup of the candidate pool. In Case 3, I illustrated how divergent responses to the moderating incentive of RCV may have swung a hotly-contested election. In these studies, I identify additional variables that I could not include in my previous analyses and detail the challenges of determining exactly how candidates adjust their messaging strategy during an election.

Each of the three cases were subject to the same rules of electoral administration; Maine has not made alterations to how RCV was conducted for federal elections since implementing RCV in 2018. Furthermore, the electorate was similar across the three cases, with voters in Case 3 representing a subset of the shared voters pools in Case 1 and Case 2. Thus, the findings in this chapter provide support for the theory that the influence of the moderating effect of RCV on candidate strategy is conditioned by candidate-level factors, which were responsible for the variation between cases.

These cases also illustrate the complexity of incumbent effects, which have been treated as simple and universal in the negative campaigning literature. My quantitative findings and the broader literature indicate that incumbents should moderate to a greater degree but that races with incumbents are more negative on average. Case 1 largely followed this formula. However, in Case 2, the overall tone of campaign rhetoric was less negative than average despite the presence of an incumbent who went out of her way to be positive. In Case 3, the incumbent did not moderate despite the clear incentive to do so in a competitive race, a decision that may have cost him the election. While incumbency does provide an advantage, incumbents do not all campaign from the same playbook. Co-partisans running in similar districts may employ different strategies, and the use of negative campaigning may vary between ostensibly similar circumstances.

The clarity offered by my case studies and the lingering questions that they raise offer many directions for future research. Researchers who take the time to gather additional data are rewarded with better-defined variables and a more complete picture of the race. Candidate strategy is dependent on a confluence of circumstances at both the candidate- and electoral-level, future studies that seek to model strategy would do well to conduct additional case studies. Such case studies should make use of as much data as possible, including campaign finance data and standardized datasets of advertisements or campaign communications like those offered by the WMP. To model how strategy changes over time, scholars could track individual advertisement purchases using campaign finance data and compare additional dimensions of advertisements, such as appeals to specific emotions or mentions of specific politicians. Finally, scholars should incorporate qualitative data to fill in the gaps left by quantitative analyses. Conducting interviews

of former candidates or campaign staff and asking how RCV impacted their messaging (if at all) is one method that could help develop more accurate models of candidate behavior and strategy.

Concluding Material

Ever since San Francisco implemented RCV in 2004, scholars, journalists, and politicians have debated the costs and benefits of implementing ranked-choice voting. Despite the best efforts of researchers, RCV scholarship has not kept pace with political developments. Every election cycle, more districts have adopted the instant runoff form of RCV, altering the strategic incentives presented to candidates and, by extension, the experience of the electorate.

Past scholarship has provided some evidence to support one of the main claims made by RCV proponents: RCV elections exhibit less negative campaigning than plurality elections. The findings of this literature have been limited by the availability of relevant data and the lack of a robust theoretical model that explains *why* candidates adjust their use of negative campaigning. This thesis solves both problems by elaborating on past theories about the "moderating incentive" that the mechanics of RCV creates and by providing the most comprehensive dataset on RCV elections in the United States to date.

By applying sentiment analysis to my novel dataset, I demonstrate that the extent to which candidates engage in rhetorical moderation under RCV is conditional on differences present at the candidate-level. I show that female candidates use 10-14 percentage points less negative campaigning than their male counterparts, and that incumbents are less negative than challengers by 14 percentage points. Further, I show that races with an incumbent are 10-12 percentage points more negative on average than open-seat races, depending on the level of restriction and model specifications. By subjecting additional null results to careful scrutiny, I present evidence that the moderating incentive of RCV is robust across electoral and administrative contexts. Robustness, i.e., a lack of conditionality due to election-level features, is

critical for the future of RCV implementation given the decentralized nature of electoral administration in the United States.

However, my findings should be interpreted with some caution, and not just because of quantitative or model-related limitations. By using a qualitative case-study approach on a restricted set of federal elections, I show that candidate messaging strategy depends on additional factors not accounted for in my models. At the individual level, candidates sometimes behave counter to expectations even when my theoretical model predicts that they are most likely to moderate their use of negative campaigning. Furthermore, I illustrate the difficulty of drawing definitive conclusions about the most consequential set of RCV elections (i.e., those at the federal level) given that very few have been conducted in the United States.

My findings suggest several directions for future research. While I studied the effects of RCV implementation on candidate strategy, there are several other avenues—and untested claims made by RCV advocates—that merit further investigation. If research on ranked-choice voting in the United States is to continue, scholars must address the problem of data availability. Collecting even basic information about candidates and electoral contexts of all major RCV elections in the U.S. required enough effort to warrant an entire chapter of this thesis. The amount of original data collection required to study RCV elections is sufficient to deter many researchers and presents pitfalls for those that choose to engage in data collection. There have been far more than 34 RCV elections in the U.S. in the 21st century, and there will be more to come in future years. Additional investigation into the interaction between the reduced negativity of incumbents and the increased negativity of their challengers in any circumstances, whether RCV or plurality, is also a necessity.

Scholars must also clarify and isolate other effects of RCV on candidates and campaigns while accounting for contextual moderators. This study deals with the moderating effect of RCV on rhetoric, but research suggests RCV may also affect voter turnout and ideological stances taken by candidates. Investigating these potential interactions requires large-*n* comparative work between RCV and non-RCV states and municipalities, which have not been done at scale. The data collection component of this thesis should serve as an indication that such efforts are both possible and worth the effort.

Research into the effects of RCV should not stop at its effects on candidates and the electorate immediately after implementation. As states and municipalities continue to use RCV, similar between-election studies can determine the persistent effects of RCV over time. Once the effects of RCV are clear, policy researchers can determine whether implementing RCV is a worthwhile endeavor. Efforts to implement RCV have largely consisted of citizen-run ballot referendums, which are relatively costly forms of activism. RCV appears to alter candidate behavior only slightly, but in an age of razor-thin polling margins in prominent races, small effects may have an outsize impact on candidate rhetoric, voter behavior, and electoral outcomes. The normative value of implementing RCV remains an open question, especially in an era of administrative democratic erosion at the state level. Other electoral reforms may accomplish the goals of RCV advocates with less effort or fewer side effects, a question that demands empirical investigation.

As a final note, American researchers, *especially* scholars of American politics, must situate their research in a broader international context. Future theoretical models should attempt to engage with cross-national findings in the electoral systems, preference voting, and ranked-choice voting literature. Political scientists who study Australia (Reilly 2021; Reilly 2018),

Belgium (André et al. 2012), Scotland (Clark 2021), and broader cross-national systems (Farrell and McAllister 1996) have made extensive theoretical and empirical contributions to the study of RCV and preference-based voting. To understand the generalized effects of RCVand the factors that condition them, scholars must look beyond the United States even as they study domestic elections. Creating a broader model of how and why ranked-choice voting functions is clearly a worthwhile goal. The first steps made by this thesis serve as an indication that doing so is both achievable and important for the study of political communication and electoral systems.

Bibliography

"113th/114th Congressional District Urban-Rural Characteristics." 2013. *ProximityOne* [blog]. http://proximityone.com/cd113_2010_ur.htm

Abadie, Alberto, Athey, Susan, Imbens, Guido W., and Jeffrey M. Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics* 138(1). https://doi.org/10.1093/qje/qjac038

Adelman, Lynn. 2021. "A New Stage in the Struggle for Voting Rights." *University of Arkansas at Little Rock Law Review* 43(4): 477–98. http://dx.doi.org/10.2139/ssrn.3908609

Alaska Division of Elections. 2022. "Alaska Better Elections Implementation." *Alaska Division of Elections* [website]. https://www.elections.alaska.gov/RCV.php

Alvarez, R. Michael, Thad E. Hall, and Ines Levin. 2018. "Low-Information Voting: Evidence From Instant-Runoff Elections." *American Politics Research* 46(6): 1012–38. https://doi.org/10.1177/1532673X18759643

André, Aubrey, Wauters, Bram, & Jean-Benoit Pilet. 2012. "It's Not Only About Lists: Explaining Preference Voting in Belgium." *Journal of Elections, Public Opinion and Parties*, 22 (3), 293-313, DOI: 10.1080/17457289.2012.692374

Ansolabehere, Stephen, and Shanto Iyengar. 1995. *Going Negative*: *How Political Advertisements Shrink and Polarize the Electorate.* Free Press.

Ansolabehere, Stephen, Shanto Iyengar, Adam Simon, and Nicholas Valentino. 1994. "Does Attack Advertising Demobilize the Electorate?" *American Political Science Review* 88 (4): 829–38. DOI: 10.2307/2082710.

Armstrong, Katherine J. 2019. "Ranked Choice Voting in Maine." *William and Flora Hewlett Foundation*. https://hewlett.org/wpcontent/uploads/2019/09/RCV-in-Maine-final-for-webposting.pdf.

Auter, Zachary, and Jeffrey Fine. 2016. "Negative Campaigning in the Social Media Age: Attack Advertising on Facebook." *Political Behavior* 38: 999–1020. https://doi.org/10.1007/s11109-016-9346-8

Avis, Eric, Ferraz, Claudio, Finan, Frederico, and Carlos Varjão. 2022. "Money and Politics: The Effects of Campaign Spending Limits on Political Entry and Competition." *American Economic Journal: Applied Economics* 14(4): 167–99. DOI: 10.1257/app.20200296

Baber v. Dunlap, 376 F. Supp. 3d 125 (D. Me. 2018)

Bantum, Erin O'Carroll, and Jason E. Owen. 2009. "Evaluating the Validity of Computerized Content Analysis Programs for Identification of Emotional Expression in Cancer Narratives." *Psychological Assessment* 21: 79–88. DOI: 10.1037/a0014643

Banwart, Mary Christine, and Mitchell S. McKinney. 2005. "A Gendered Influence in Campaign Debates? Analysis of Mixed-gender United States Senate and Gubernatorial Debates." *Communication Studies* 56 (4). https://doi.org/10.1080/10510970500319443

Baranowski, Paweł, Simon Kruschinski, Uta Russmann, Jörg Haßler, Melanie Magin, Bene Márton, Andrea Ceron, Daniel Jackson, and Darren Lilleker. 2022. "Patterns of Negative Campaigning during the 2019 European Election: Political Parties' Facebook Posts and Users' Sharing Behaviour across Twelve Countries." *Journal of Information Technology & Politics*: 1–18. https://doi.org/10.1080/19331681.2022.2115598

Barber, Kathleen L. 2000. *A Right To Representation: Proportional Election Systems for the Twenty-First Century.* Ohio State University Press.

Bauer, Nichole M., and Martina Santia. 2023. "Gendered Times: How Gendered Contexts Shape Campaign Messages of Female Candidates" *Journal of Communication* 353—73. https://academic.oup.com/joc/advance-article/doi/10.1093/joc/jqac052/6982618

Bekafigo, Marija Anna, and Allison Clark Pingley. 2015. "Tweeting Negative: Determinants of Negative Campaigning in the 2011 Gubernatorial Elections." *International Journal of E-Politics* 6 (1): 30–41. https://doi.org/10.4018/IJEP.2015010103.

Benoit, Kenneth, Watanabe, Kohei, Wang, Haiyan, Nulty, Paul, Obeng, Adam, Müller, Stefan, and Akitaka Matsuo. 2018. "quanteda: An R package for the quantitative analysis of textual data." *Journal of Open Source Software* 3(30): 774. doi:10.21105/joss.00774

Benoit, William L., Wen, Wei-Chun, and Yu, Tzu-Hsiang. 2007. "A functional analysis of 2004 Taiwanese political debates." *Asian Journal of Communication* 17 (1): 24-39. https://doi.org/10.1080/01292980601114521

Benoit, William L. 2014. *A functional analysis of political television advertisements.* Lexington Books.

Bernhardt, Dan, and Meenakshi Ghosh. 2020. "Positive and Negative Campaigning in Primary and General Elections." *Games and Economic Behavior* 119: 98–104. https://doi.org/10.1016/j.geb.2019.10.011.

Bode, Leticia, and Kajsa E. Dalrymple. 2016. "Politics in 140 Characters or Less: Campaign Communication, Network Interaction, and Political Participation on Twitter." *Journal of Political Marketing* 15(4): 311–32. https://doi.org/10.1080/15377857.2014.959686

Bolton, Alexander. 2020. "Polls show Collins displaces McConnell as most unpopular senator." *The Hill*. https://thehill.com/homenews/senate/478702-poll-shows-collins-displaces-mcconnell-as-most-unpopular-senator/

Bryan, Amanda C., and Eve M. Ringsmuth. 2016. "Jeremiad or Weapon of Words? The Power of Emotive Language in Supreme Court Dissents." *Journal of Law and Courts* 4 (1): 159–85. doi:10.1086/684788

Burden, Barry. 2002. "When Bad Press Is Good News: The Surprising Benefits of Negative Campaign Coverage." *Harvard International Journal of Press/Politics,* 7(3), 76–89. https://doi.org/10.1177/1081180X0200700305

Burden, Berry. 2009. "WiscAds 1996—2004 & 2008." University of Wisconsin Madison [producer], Wisconsin Advertising Project. https://elections.wisc.edu/wisconsin-advertising-project/.

Callender, James Thomson. 1800. *The Prospect Before Us*.

Cameron, A. Colin, Gelbach, Jonah B., and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, 90(3). doi: https://doi.org/10.1162/rest.90.3.414

Caren, Neal. 2007. "Big City, Big Turnout? Electoral Participation in American Cities." *Journal of Urban Affairs*, 29: 31-46. https://doi.org/10.1111/j.1467-9906.2007.00321.x

Carman, Erin, and Vanessa Glushefski. 2019. "Ranked Choice Voting Now: A Shift toward a Better Democracy." *Social Justice & Equity Law Journal*, 3 (2): 36-74. https://digitalcommons.daemen.edu/cgi/viewcontent.cgi?article=1194&context=faculty_scholar

Cauterucci, Christina. 2020. "How Maine Turned on Susan Collins." *Slate*. https://slate.com/news-and-politics/2020/09/maine-turned-on-susan-collins.html

Cellini, Nicola, Natale Canale, Giovanna Mioni, and Sebastiano Costa. 2020. "Changes in Sleep Pattern, Sense of Time and Digital Media Use during COVID-19 Lockdown in Italy." *Journal of Sleep Research* 29(4). https://doi.org/10.1111/jsr.13074

Ceron, Andrea, and Giovanna d'Adda. 2016. "E-Campaigning on Twitter: The Effectiveness of Distributive Promises and Negative Campaign in the 2013 Italian Election." *New Media & Society* 18(9): 1935–55. https://doi.org/10.1177/1461444815571915

Cillizza, Chris. 2020. "7 of the 10 most expensive Senate races ever are happening right now." *CNN*. https://www.cnn.com/2020/10/15/politics/senate-races-fundraising-2020-election/index.html

City of Minneapolis. 2022. "How We Count RCV Ballots." *City of Minneapolis* [website]. https://vote.minneapolismn.gov/ranked-choice-voting/how-we-count-rcv-ballots/

Clark, Alistair. 2021. “The Effects of Electoral Reform on Party Campaigns, Voters and Party Systems at the Local Level: From Single Member Plurality to the Single Transferable Vote in Scotland.” *Local Government Studies* 47(1): 79–99.

Clark, Jesse. 2020. “Rank Deficiency? Analyzing the Costs and Benefits of Single-Winner Ranked-Choice Voting.” *SSRN Scholarly Paper*. https://doi.org/10.2139/ssrn.3703197

Coll, J. A. (2021). “Demographic disparities using ranked-choice voting? Ranking difficulty, under-voting, and the 2020 democratic primary.” *Politics and Governance* 9 (2): 293-305. https://doi.org/10.17645/pag.v9i2.3913

Collins, Steve. 2018. “Poliquin indicates he might challenge ranked-choice results if he doesn’t win.” *Lewiston Sun Journal*. https://www.centralmaine.com/2018/10/22/poliquin-indicates-he-might-challenge-election-results-if-he-doesnt-win/

“Collins, Susan Margaret.” 2020. *Voteview* [website]. https://voteview.com/person/49703/susan-margaret-collins

“County Commissioner Districts: Unified Proposal.” 2021. *Maine State Legislature* [website], September 20. https://legislature.maine.gov/apportionment-commission/available-maps-92121

Cox, Gary W., Fiva, Jon H. ,and Daniel M. Smith. 2020. “Measuring the Competitiveness of Elections.” *Political Analysis* 28 (2): 168–85. https://doi.org/10.1017/pan.2019.28

Crabtree, Charles, Golder, Matt, Gschwend, Thomas, and Indridi H. Indriđason. 2020. “It is not only what you say, it is also how you say it: The strategic use of campaign sentiment.” *The Journal of Politics*, *82*(3), 1044-1060. https://doi.org/10.1086/707613

Cronert, Axel, and Pär Nyman. 2021. “A General Approach to Measuring Electoral Competitiveness for Parties and Governments.” *Political Analysis* 29 (3): 337–55. https://doi.org/10.1017/pan.2020.30

Davidov, Dmitry, Tsur, Oren, and Rappoport, Ari. 2010. “Enhanced Sentiment Learning Using Twitter Hashtags and Smileys.” *ICCL-10.* https://aclanthology.org/C10-2028

De Choudhury, Munmun, Counts, Scott, and Eric Horvitz. 2013. “Predicting Postpartum Changes in Emotion and Behavior via Social Media.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 3267-76. https://doi.org/10.1145/2470654.2466447

Donovan, Todd, Caroline Tolbert, and Kellen Gracey. 2016. “Campaign Civility under Preferential and Plurality Voting.” *Electoral Studies* 42: 157–63. https://doi.org/10.1016/j.electstud.2016.02.009.

Druckman, James N., Kifer, Martin J., and Michael Parkin. 2020. "Campaign Rhetoric and the Incumbency Advantage." *American Politics Research* 48 (1): 22-43. https://doi.org/10.1177/1532673X18822314

Elbagir, Shihab, and Jing Yang. 2019. "Twitter sentiment analysis using natural language toolkit and VADER sentiment". *Proceedings of the international multiconference of engineers and computer scientists* 122(16). https://www.iaeng.org/publication/IMECS2019/IMECS2019_pp12-16.pdf

FairVote. 2022. "Where is Ranked Choice Voting Used?" *FairVote* [website], September. https://fairvote.org/our-reforms/ranked-choice-voting-information/#where-is-ranked-choice-voting-used

FairVote. 2023. "Ranked Choice Voting." *FairVote* [website]. https://fairvote.org/our-reforms/ranked-choice-voting/

Farrell, David M., and Ian Mcallister. 2006. "Voter Satisfaction and Electoral Systems: Does Preferential Voting in Candidate-Centred Systems Make a Difference?" *European Journal of Political Research* 45(5): 723–49.

Finkel, Steven .E., and John G. Geer. 1998. "A Spot Check Casting Doubt on Demobilizing Effect of Attack Advertising." *American Journal of Political Science* 42(2): 573–595. https://doi.org/10.2307/2991771

Fowler, Erika Franklin, Franz, Michael M., and Travis N. T. N. 2016. *Political advertising in the United States*. 1st ed. Westview Press.

Fowler, Erika Franklin, Michael M. Franz, and Travis N. Ridout. 2023. Political Ad Airings Data, 2006 & 2010—2020, Wesleyan University (producer), Wesleyan Media Project. https://mediaproject.wesleyan.edu/dataaccess/

Fowler, Erika Franklin, Franz, Michael M., and Travis N. Ridout. 2021(a). "Buying and Targeting Political Advertising". In *Political Advertising in the United States* 2nd ed. Routledge.

Fowlerb, Erika Franklin, Franz, Michael M., Martin, Gregory, J., Peskowitz, Zachary, and Travis N. Ridout. 2021(b). "Political Advertising Online and Offline." *American Political Science Review* 115(1): 130–49. doi:10.1017/S0003055420000696

Fowler, Erika Franklin, Michael M. Franz, Travis N. Ridout, Laura M. Baum, and Colleen Bogucki. 2023. "2020 Political TV Advertising." Version 1.0 [dataset]. Middletown, CT: The Wesleyan Media Project, Wesleyan University.

Fowler, Erika Franklin, Michael M. Franz, Travis N. Ridout, Laura M. Baum, and Colleen Bogucki. 2023. "2018 Political TV Advertising." Version 1.0 [dataset]. Middletown, CT: The Wesleyan Media Project, Wesleyan University.

Fuchs, C. 2021. "How did Donald Trump incite a coup attempt?" *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 19(1), 246-251. https://doi.org/10.31269/triplec.v19i1.1239

Galasso, Vincenzo, Tommaso Nannicini, and Salvatore Nunnari. 2021. "Positive Spillovers from Negative Campaigning." *American Journal of Political Science.* https://doi.org/10.1111/ajps.12610

Gastil, John. 2022. "Can Voters Trust Each Other? Designing an Online Model for Non-Interactive Issue Deliberation to Produce a Trustworthy Voting Aid in Direct Democratic Elections." Paper Presented at the American Political Science Association Annual Meeting & Exhibition in Montreal, Canada.

Geer, John G. *In defense of negativity: Attack ads in presidential campaigns*. Chicago, IL; University of Chicago Press, 2008.

Gonzales, Nathan L. 2023. "Is the competitive House battleground due to shrink?" *Roll Call* [blog], January 25. https://rollcall.com/2023/01/25/is-the-competitive-house-battleground-due-to-shrink/

Gross, Justin H., and Kaylee T. Johnson. 2016. "Twitter Taunts and Tirades: Negative Campaigning in the Age of Trump." *Political Science & Politics* 49 (4): 748–54. https://doi.org/10.1017/S1049096516001700.

Gutiérrez, Manuel, Simmons, Alan J. and John Transue. 2022. "Ranked-Choice Voting and Democratic Attitudes." *American Politics Research*. DOI: 1532673X221109519

Hajnal, Zoltan L., & Paul G. Lewis. 2003. "Municipal Institutions and Voter Turnout in Local Elections." *Urban Affairs Review*, 38(5), 645–668. https://doi.org/10.1177/1078087403038005002

Hansen, Kasper M., and Kosiara-Pedersen, Karina. 2017. "How campaigns polarize the electorate: Political polarization as an effect of the minimal effect theory within a multi-party system." *Party Politics*, 23 (3): 181-192. https://doi.org/10.1177/1354068815593453

Haselmayer, Martin. 2019. "Negative Campaigning and Its Consequences: A Review and a Look Ahead." *French Politics* 17 (3): 355–72. https://doi.org/10.1057/s41253-019-00084-8

Hassell, Hans J.G. 2021. "Desperate Times Call for Desperate Measures: Electoral Competitiveness, Poll Position, and Campaign Negativity." *Political Behavior* 43: 1137–59. https://doi.org/10.1007/s11109-019-09588-z

Hemsley, Jeff, Stromer-Galley, Jennifer, Semaan, Bryan, and Sikana Tanupabrungsun. 2018. "Tweeting to the Target: Candidates' Use of Strategic Messages and @Mentions on

Twitter." *Journal of Information Technology & Politics* 15(1): 3–18. https://doi.org/10.1080/19331681.2017.1338634

Hirano, Shigeo, Snyder, James M., Ansolabehere, Stephen D., and John Mark Hansen. 2010. "Primary elections and partisan polarization in the U.S. Congress." *Quarterly Journal of Political Science* 5 (2): 169-191. https://doi.org/10.1561/100.00008052

Holbrook, Thomas M., and Aaron C. Weinschenk. 2014. "Campaigns, Mobilization, and Turnout in Mayoral Elections." *Political Research Quarterly* 67 (1): 42–55. https://doi.org/10.1177/1065912913494018

Hopmann, David Nicolas, Rens Vliegenthart, and Jürgen Maier. 2018. "The Effects of Tone, Focus, and Incivility in Election Debates." *Journal of Elections, Public Opinion and Parties* 28(3): 283–306. https://doi.org/10.1080/17457289.2017.1394310

Hosch-Dayican, Bengü, Chintan Amrit, Kees Aarts, and Adrie Dassen. 2016. "How Do Online Citizens Persuade Fellow Voters? Using Twitter During the 2012 Dutch Parliamentary Election Campaign." *Social Science Computer Review* 34 (2): 135–52. https://doi.org/10.1177/0894439314558200

"How Maine Voted: Governors' races 1990-2014". 2015. *The Portland Press-Herald* [interactive tool]. https://www.pressherald.com/interactive/maine-voted-governors-races-1990-2010/

Huang, Francis L. (2018). "Using Cluster Bootstrapping to Analyze Nested Data With a Few Clusters." *Educational and Psychological Measurement*, 78(2): 297–318. https://doi.org/10.1177/0013164416678980

Humphry, Debbie. 2020. "Campaigning in the Time of Coronavirus." *City* 24(3–4): 423–30. https://doi.org/10.1080/13604813.2020.1791551

Hutto, C., and Gilbert, Eric. 2014. "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text." *Proceedings of the International AAAI Conference on Web and Social Media*, 8 (1), 216-225. https://doi.org/10.1609/icwsm.v8i1.14550

Iyengar, Shanto, and Sean J. Westwood. 2015. "Fear and Loathing across Party Lines: New Evidence on Group Polarization." *American Journal of Political Science* 59 (3): 690–707. https://doi.org/10.1111/ajps.12152.

Jacobson, Gary C. 2019. "Extreme Referendum: Donald Trump and the 2018 Midterm Elections." *Political Science Quarterly*, 134(1): 9–38. https://doi.org/10.1002/polq.12866

John, Sarah, and Andrew Douglas. 2017. "Candidate Civility and Voter Engagement in Seven Cities with Ranked Choice Voting." *National Civic Review* 106 (1): 25–29. https://www.jstor.org/stable/10.1002/naticivirevi.106.1.0025

John, Sarah, Smith, Haley, and Elizabeth Zack. 2018. "The Alternative Vote: Do Changes in Single-Member Voting Systems Affect Descriptive Representation of Women and Minorities?" *Electoral Studies* 54: 90–102. https://doi.org/10.1016/j.electstud.2018.05.009

Johnson, Dennis W. 2007. *No Place for Amateurs: How Political Consultants Are Reshaping American Democracy*. 2nd ed. New York: Routledge.

Juelich, Courtney L., and Joseph A. Coll. 2021. "Ranked Choice Voting and Youth Voter Turnout: The Roles of Campaign Civility and Candidate Contact." *Politics and Governance* 9 (2): 319–31. https://doi.org/10.17645/pag.v9i2.3914.

Kahn, Jeffrey H., Tobin, Renée M., Massey, Audra E. , and Jennifer A. Anderson. 2007. "Measuring Emotional Expression with the Linguistic Inquiry and Word Count." *The American Journal of Psychology*. http://www.jstor.org/stable/20445398

Kahn, Kim F. 2016. "Gender Differences in Campaign Messages: The Political Advertisements of Men and Women Candidates for U.S. Senate." *Political Research Quarterly*, 46 (3), 481–502. https://doi.org/10.1177/106591299304600303

Kalsnes, Bente, Krumsvik, Arne H., and Tanja Storsul. 2014. "Social Media as a Political Backchannel: Twitter Use during Televised Election Debates in Norway." *Aslib Journal of Information Management*. DOI: 10.1108/AJIM-09-2013-0093

Karlsen, Rune, and Bernard Enjolras. 2016. "Styles of Social Media Campaigning and Influence in a Hybrid Political Communication System: Linking Candidate Survey Data with Twitter Data." *The International Journal of Press/Politics* 21(3): 338–57. https://doi.org/10.1177/1940161216645335

Keller, Betty. 2017. "Pros and Cons of Instant Runoff (Ranked Choice) Voting". *League of Women Voters Vermont* [blog]. https://my.lwv.org/vermont/article/pros-and-cons-instant-runoff-ranked-choice-voting.

Keller, Tobias R., and Katharina Kleinen-von Königslöw. 2018. "Followers, Spread the Message! Predicting the Success of Swiss Politicians on Facebook and Twitter." *Social Media + Society* 4(1). https://doi.org/10.1177/2056305118765733

Kennedy, Courtney, Blumenthal, Mark, Clement, Scott, Clinton, Joshua D., Durang, Claire, Franklin, Charles, McGeeney, Kyley, Miringoff, Lee, Olson, Kristen, Rivers, Douglas, Saad, Lydia, Witt, G Evans, and Christopher Wlezien. 2018. "An Evaluation of the 2016 Election Polls in the United States." *Public Opinion Quarterly* 82(1): 1–33. https://doi.org/10.1093/poq/nfx047

Kennedy, Ian, Wack, Morgan, Beers, Andrew, Schafer, Joseph S., Garcia-Camargo, Isabella, Spiro, Emma S., and Kate Starbird. 2022. "Repeat Spreaders and Election Delegitimization: A Comprehensive Dataset of Misinformation Tweets from the 2020

U.S. Election." *Journal of Quantitative Description: Digital Media* 2. https://doi.org/10.51685/jqd.2022.013

Kimball, David, and Joseph Anthony. 2018. "Ranked Choice Voting: A Different Way of Casting and Counting Votes." In *Changing How America Votes*, ed. Todd Donovan, 100—112. Lanham, MD; Rowman & Littlefield.

"King, Angus Stanley, Jr." 2018. *Voteview*, [website]. https://voteview.com/person/41300/angus-stanley-king-jr

Kropf, Martha. 2021. "Using Campaign Communications to Analyze Civility in Ranked Choice Voting Elections." *Politics and Governance* 9 (2): 280–92. https://doi.org/10.17645/pag.v9i2.4293

Lau, Richard R. 1985. "Two Explanations for Negativity Effects in Political Behavior." American Journal of Political Science 29 (1): 119–38. https://doi.org/10.2307/2111215.

Lau, Richard R., and Gerald M. Pomper. 2004. "Who Uses Negative Campaigning?" In *Negative Campaigning: An Analysis of U.S. Senate Elections*. Rowman & Littlefield.

Lau, Richard R., and Pomper, Gerald M. 2002. "Effectiveness of Negative Campaigning in U.S. Senate Elections." *American Journal of Political Science*. 46 (1), 47-66. https://www.jstor.org/stable/3088414

Lau, Richard R., and Rovner, Ivy Brown. 2009. "Negative campaigning." *Annual Review of Political Science* 12: 285-306. DOI: 10.1146/annurev.polisci.10.071905.101448

Lau, Richard R., Lee Sigelman, Caroline Heldman, and Paul Babbitt. 1999. "The Effects of Negative Political Advertisements: A Meta-Analytic Assessment." *American Political Science Review* 93 (4): 851–75. DOI: 10.2307/2586117.

League of Women Voters of Maine. 2020. "Ranked Choice Voting." *League of Women Voters of Maine* [website]. https://www.lwvme.org/RCV

Lee, Jayeon, and Weiai Xu. 2018. "The More Attacks, the More Retweets: Trump's and Clinton's Agenda Setting on Twitter." *Public Relations Review* 44(2): 201–13. https://doi.org/10.1016/j.pubrev.2017.10.002

Leifeld, Philip. 2013. "texreg: Conversion of Statistical Model Output in R to LaTeX and HTML Tables." *Journal of Statistical Software* 55(8): 1–24. http://dx.doi.org/10.18637/jss.v055.i08

Li, Michael, and Chris Leaverton. 2022. "Gerrymandering Competitive Districts to Near Extinction." *Brennan Center for Justice.* https://www.brennancenter.org/our-work/analysis-opinion/gerrymandering-competitive-districts-near-extinction.

Liu, Bing. 2012. "Sentiment analysis and opinion mining." In *Synthesis lectures on human language technologies,* 5 (2), 1-167. Morgan and Claypool Publishers.

Maier, Jürgen, and Carolin Jansen. 2017. "When Do Candidates Attack in Election Campaigns? Exploring the Determinants of Negative Candidate Messages in German Televised Debates." *Party Politics* 23(5): 549–59. https://doi.org/10.1177/1354068815610966

Maine State Legislature. 2022. "Ranked Choice Voting in Maine." *Maine State Legislature Law Library* [website]. https://legislature.maine.gov/lawlibrary/ranked-choice-voting-in-maine/9509

Maizland, Lindsay. 2020. "How Countries Are Holding Elections During the COVID-19 Pandemic." *Council on Foreign Relations.* http://www.jstor.org/stable/resrep29836

Marien, Sofie, Ine Goovaerts, and Stephen Elstub. 2020. "Deliberative Qualities in Televised Election Debates: The Influence of the Electoral System and Populism." *West European Politics* 43(6): 1262–84. https://doi.org/10.1080/01402382.2019.1651139

Mascaro, Christopher, and Sean Patrick Goggins. 2012. "Twitter as Virtual Town Square: Citizen Engagement During a Nationally Televised Republican Primary Debate." Presented at the Annual Meeting and Exhibition of the American Political Science Association; New Orleans, Louisiana. https://papers.ssrn.com/abstract=2108682

Mattes, Kyle, and David P. Redlawsk. 2015. *The Positive Case for Negative Campaigning.* Chicago, IL; University of Chicago Press.

Matthew, Karen, and Deepthi Hajela. 2021. "Error Mars Vote Count in NYC Mayoral Primary." *The Associated Press.* https://apnews.com/article/eric-adams-lead-shrinks-nyc-democratic-mayor-primary-0c92450b5dbb57018e2f8e15d8471a97

McCannon, Bryan C., 2022. "Ranked Choice Voting in Mayoral Elections". *SSRN,* http://dx.doi.org/10.2139/ssrn.4100734

McDaniel, Jason A. 2019. "Electoral Rules and Voter Turnout in Mayoral Elections: An Analysis of RCV." Paper presented at the Election Systems, Reform, and Administration Conference in Philadelphia, Pennsylvania.

McDaniel, Jason. 2018. "Does More Choice Lead to Reduced Racially Polarized Voting? Assessing the Impact of RCV in Mayoral Elections." *California Journal of Politics and Policy* 10(2). https://escholarship.org/uc/item/2gm5854x

McGing, Claire. 2013. "The Single Transferable Vote and Women's Representation in Ireland." *Irish Political Studies* 28(3): 322–40. https://doi.org/10.1080/07907184.2013.818536

McGinn, Eamon. 2020. "Rating rankings: Effect of instant run-off voting on participation and civility." Unpublished manuscript. http://eamonmcginn.com/papers/IRV_in_Minneapolis.pdf

Mueller, Samuel David, and Marius Saeltzer. 2022. "Twitter Made Me Do It! Twitter's Tonal Platform Incentive and Its Effect on Online Campaigning." *Information, Communication & Society* 25 (9): 1247–72. https://doi.org/10.1080/1369118X.2020.1850841.

Murthy, Dhiraj. 2015. "Twitter and Elections: Are Tweets, Predictive, Reactive, or a Form of Buzz?" *Information, Communication & Society* 18(7): 816–31. https://doi.org/10.1080/1369118X.2015.1006659

Nai, Alessandro. 2020. "Going Negative, Worldwide: Towards a General Understanding of Determinants and Targets of Negative Campaigning." *Government and Opposition* 55 (3): 430–55. https://doi.org/10.1017/gov.2018.32.

Neely, Francis, and Corey Cook. 2008. "Whose Votes Count?: Undervotes, Overvotes, and Ranking in San Francisco's Instant-Runoff Elections." *American Politics Research* 36(4): 530–54. https://doi.org/10.1177/1532673X08318110

Nemitz, Bill. 2018. "On gun laws, Jared Golden thinks just like most Mainers." *The Portland Press-Herald*. https://www.pressherald.com/2018/09/17/on-gun-laws-jared-golden-thinks-just-like-most-mainers/

"New York Times Upshot/Siena College: October 2018 Maine 02 Poll." 2018. *The New York Times*.

"New York Times Upshot/Siena College: September 2018 Maine 02 Poll." 2018. *The New York Times.*

Nielson, Lindsay. 2017. "Ranked Choice Voting and Attitudes toward Democracy in the United States: Results from a Survey Experiment." *Politics & Policy* 45(4): 535–70. https://doi.org/10.1111/polp.12212

Panagopoulos, Costas. 2004. "Boy Talk/Girl Talk." *Women & Politics* 26 (3–4): 131–55. https://doi.org/10.2307/448944

Parker, Kim, Horowitz, Juliana M., Brown, Anna, Fry, Richard, Cohn, D'Vera, and Ruth Igielnik. 2018. "Urban, suburban and rural residents' views on key social and political issues." In "What Unites and Divides Urban, Suburban and Rural Communities.*" Pew Research Center*. https://www.pewresearch.org/social-trends/2018/05/22/urban-suburban-and-rural-residents-views-on-key-social-and-political-issues/

Parmelee, John H., and Shannon L. Bichard . 2011. *Politics and the Twitter Revolution: How Tweets Influence the Relationship between Political Leaders and the Public.* Lexington Books.

Pennebaker, James W, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The Development and Psychometric Properties of LIWC2015.* Austin, TX: University of Texas at Austin.

Peterson, David A. M., and Paul A. Djupe. 2005. "When Primary Campaigns Go Negative: The Determinants of Campaign Negativity." *Political Research Quarterly* 58 (1): 45–54. https://doi.org/10.1177/106591290505800104.

Poole, Keith T. 2005. *Spatial models of parliamentary voting*. Cambridge University Press.

Preoţiuc-Pietro, Daniel, Ye Liu, Daniel Hopkins, and Lyle Ungar. 2017. "Beyond Binary Labels: Political Ideology Prediction of Twitter Users." In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*: 729–40. https://aclanthology.org/P17-1068.

"Quinnipiac University Maine Poll." 2020. 31117688.00008. *Quinnipiac University Polling Institute*.

Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58(4): 1083–91. https://doi.org/10.1111/ajps.12102

Rainie, Lee, Smith, Aaron, Schlozman, Kay Lehman, Brady, Henry, and Sidney Verba. 2012. "Social Media and Political Engagement." *Pew Research Center*. https://www.pewresearch.org/internet/2012/10/19/social-media-and-political-engagement/

Ramteke, Jyoti, Shah, Samarth, Godhia, Darshan, and Aadil Shaikh. 2016. "Election result prediction using Twitter sentiment analysis." Paper presented at the International conference on inventive computation technologies 1(1).

Rank the Vote. 2023. "How Ranked Choice Voting Works." *Rank the Vote* [website]. https://rankthevote.us/learn/

Ranked Choice Voting Resource Center. 2022. "History of Ranked-Choice Voting." *Ranked Choice Voting Resource Center* [website]. https://www.rcvresources.org/history-of-rcv

Reilly, Benjamin. 2018. "Centripetalism and Electoral Moderation in Established Democracies", *Nationalism and Ethnic Politics*, 24:2, 201-221, DOI: 10.1080/13537113.2018.1457827

Reilly, Benjamin. 2021. "Ranked Choice Voting in Australia and America: Do Voters Follow Party Cues?" *Politics and Governance*, 9(2), https://doi.org/10.17645/pag.v9i2.3889

Reimann, Nicholas. 2023. "Twitter Boosts Character Limit to 4,000 for Twitter Blue Subscribers." *Forbes*. https://www.forbes.com/sites/nicholasreimann/2023/02/08/twitter-boosts-character-limit-to-4000-for-twitter-blue-subscribers/?sh=6d3b64ff5ab8

Ridout, Travis N., Franz, Michael, Goldstein, Kenneth M. and William J. Feltus. 2012. "Separation by Television Program: Understanding the Targeting of Political Advertising in Presidential Elections." *Political Communication* 29(1): 1–23. https://doi.org/10.1080/10584609.2011.619509

Rinker, Tyler W. 2020. “qdap: Quantitative Discourse Analysis Package.” 2.4.2, https://github.com/trinker/qdap

Robb, Denise Munro. 2022. “The Effect of Instant Runoff Voting on Democracy.” Ph.D. Thesis, University of California, Irvine. https://www.proquest.com/docview/912208018/abstract/D9F45C2ACB8D4791PQ/1.

Rozin, Paul, and Edward B. Royzman. 2001. "Negativity bias, negativity dominance, and contagion." *Personality and Social Psychology Review* 5 (4): 296-320. https://doi.org/10.1207/S15327957PSPR0504_2

Rudkowsky, Elena, Haselmayer, Martin, Wastian, Matthias, Jenny, Marcelo, Emrich, Štefan and Sedlmair, Michael. 2018. "More than bags of words: Sentiment analysis with word embeddings." *Communication Methods and Measures* 12 (2-3): 140-157. https://doi.org/10.1080/19312458.2018.1455817

Rudman, Laurie A., Moss-Racusin, Corinne A., Phelan, Julie E., and Sanne Nauts. 2012. “Status Incongruity and Backlash Effects: Defending the Gender Hierarchy Motivates Prejudice against Female Leaders.” *Journal of Experimental Social Psychology.* 48(1): 165–79. https://doi.org/10.1016/j.jesp.2011.10.008

Saedeen, Mohammad, and Naser N. AlBzour. 2022. “Donald Trump’s Denial Speeches of the 2020 United States Presidential Election’s Results: A Critical Discourse Analysis Perspective.” *Advances in Language and Literary Studies* 13(1): 32–40. https://doi.org/10.7575/aiac.alls.v.13n.1.p.32

Salama, Amir H. Y. 2022. “Cross-Textual Reconceptualisation of the Deictic Space of ‘Victory’ in Political Discourse: Donald Trump versus Joseph Biden.” *Discourse and Interaction* 15(1): 101–31. https://doi.org/10.5817/DI2022-1-101

San Francisco Department of Elections. 2023. “Ranked-Choice Voting.” *San Francisco Department of Elections* [website]. https://sfelections.sfgov.org/ranked-choice-voting

Santucci, Jack, and Chris Hughes. 2017. “History of Ranked-Choice Voting in the United States” [webinar], September 8. *Ranked Choice Voting Resource Center*. https://www.rcvresources.org/history-of-rcv-in-the-united-states.

Santucci, Jack, and Jamil Scott. 2021. “Do Ranked Ballots Stimulate Candidate Entry?” *SSRN.* https://papers.ssrn.com/abstract=3956554

Santucci, Jack. 2018. “Maine Ranked-Choice Voting as a Case of Electoral-System Change.” *Representation* 54(3): 297–311. http://dx.doi.org/10.1080/00344893.2018.1502208

Santucci, Jack. 2021. “Variants of Ranked-Choice Voting from a Strategic Perspective.” *Politics and Governance* 9. https://doi.org/10.17645/pag.v9i2.3955.

Scanlon, Jesse. 2018. “Could Maine’s new ranked-choice voting change American elections?” *The Boston Globe*, October 17. https://www.bostonglobe.com/magazine/2018/10/17/could-maine-new-ranked-choice-voting-change-american-elections/6VqNC73bQzMrPd0RSepA8L/story.html

Schleicher, David. 2007. “Why Is There No Partisan Competition in City Council Elections: The Role of Election Law.” *Journal of Law & Politics* 23(4): 419–74. https://ssrn.com/abstract=1122422

Schuerman, Ben. 2010. “Third Parties and City Council Elections.” *Xavier Journal of Politics*, 1(1): 67- 77.

“Senator Angus King.” 2018. *GovTrack* [website]. https://www.govtrack.us/congress/members/angus_king/412545

Shepherd, Michael. 2018. “Poliquin announces he is giving up legal challenge to ranked-choice voting.” *Bangor Daily News*. https://bangordailynews.com/2018/12/24/news/poliquin-announces-that-he-is-giving-up-legal-challenge-to-ranked-choice-voting/

Siegert, Ingo, Sinha, Yamini, Jokisch, Oliver, and Andreas Wendemuth. 2020. “Recognition Performance of Selected Speech Recognition APIs – A Longitudinal Study.” *In Speech and Computer, Lecture* Notes *in Computer Science*, eds. Alexey Karpov and Rodmonga Potapova. Cham: Springer International Publishing, 520–29.

Simon, Christopher A., and Raymond Tatalovich. 2022. “The Turnout Myth and Referendum Voting in the United States” *Wiley Online Library of Politics and Policy.* https://doi.org/10.1111/polp.12483

Stack, Liam. 2018. “He’s a Long-Shot Senate Candidate With a Message: ‘Capitalism Unchecked Is a Complete Disaster.’” *The New York Times*. https://www.nytimes.com/2018/10/22/us/politics/capitalism-socialism-america-election.html

Stein, Kevin A., and Benoit, William L. 2021. “A functional analysis of 2016 nonpresidential campaign tweets.” *American Behavioral Scientist*, 65 (3): 432-447.

“Suffolk University Maine Poll.” 2018. 31116584.00013. *Suffolk University Political Research Center.*

“Tabulations for Elections Held in 2018: November 6, 2018 General Election.” 2018. *Maine Bureau of Corporations, Elections & Commissions* [dataset]. Available at: https://www.maine.gov/sos/cec/elec/results/results18.html

“Tabulations for Elections Held in 2020: November 3, 2020 General Election.” 2020. *Maine Bureau of Corporations, Elections & Commissions* [dataset]. Available at: https://www.maine.gov/sos/cec/elec/results/results20.html

Tausczik, Yla R., and James W. Pennebaker. 2010. "The psychological meaning of words: LIWC and computerized text analysis methods." *Journal of language and social psychology*, 29(1), 24-54. https://doi.org/10.1177/0261927X09351676

Tumasjan, Andranik, Sprenger, Timm O., Sandner, Philipp G., and Isabell M. Welpe. 2011. "Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape." *Social Science Computer Review* 29(4): 402–18. https://doi.org/10.1177/0894439310386557

Valli, Chiara, and Alessandro Nai. 2020. "Attack Politics from Albania to Zimbabwe: A Large-Scale Comparative Study on the Drivers of Negative Campaigning." *International Political Science Review*. https://doi.org/10.1177/0192512120946410.

Van Kessel, Patrick, Widijaya, Regina, Shah, Sono, Smith, Aaron, and Adam Hughes. 2020. "The congressional social media landscape." In "Congress Soars to New Heights on Social Media." *Pew Research Center.* https://www.pewresearch.org/internet/2020/07/16/1-the-congressional-social-media-landscape/

Walter, Annemarie S., and Rens Vliegenthart. 2010. "Negative Campaigning across Different Communication Channels: Different Ball Games?" *The International Journal of Press/Politics* 15(4): 441–61. https://doi.org/10.1177/1940161210374122

Walter, Annemarie, and Cees van der Eijk. 2019. "Unintended consequences of negative campaigning: Backlash and second-preference boost effects in a multi-party context." *The British Journal of Politics and International Relations*, 21 (3): 612–629. https://doi.org/10.1177/1369148119842038

Wang, Hui H. 2021. "Speech Recorder and Translator Using Google Cloud Speech-to-Text and Translation." *Journal of IT in Asia.* https://publisher.unimas.my/ojs/index.php/JITA/article/view/2815

Warren, Elizabeth, and Raskin, Jamie. 2020. "Ranked-choice voting is a better way to vote." *The Boston Globe,* September 18. https://www.bostonglobe.com/2020/09/18/opinion/ranked-choice-voting-is-better-way-vote/

West, Darrell M. 2017. *Air wars: television advertising and social media in election campaigns, 1952-2016.* CQ Press.

"Why Portland Doesn't Have an Elected Mayor." 2010. *Maine Voices* [blog], September 27. https://www.pressherald.com/2010/09/27/why-portland-doesnt-have-an-elected-mayor_2010-09-27/

Willis, Derek. 2023. "After a Decade of Tracking Politicians' Deleted Tweets, Politwoops Is No More." *ProPublica*. https://www.propublica.org/article/politwoops-deleted-tweets-twitter-politicians-musk

Xu, Haiyang, Zhang, Hui, Han, Kun, Wang, Yun, Peng, Yiping, and Li Xiangang. 2020. "Learning Alignment for Multimodal Emotion Recognition from Speech." *Proc. Interspeech 2019*, 3569-3573. http://arxiv.org/abs/1909.056