Does Ranked-Choice Voting Reduce Racial Polarization? Evidence from Cast Vote Records*

Yuki Atsusaka† Theodore Landsman‡

Current version: January 3, 2019

Abstract

In this article, we provide the most direct test of the moderation effect of implementing ranked-choice voting (RCV) on the amount of racially polarized voting. First, we collect and analyze cast vote records where we have the known truth on individual rankings on available candidates in RCV elections. Second, drawing from the rank data literature, we explore potential clusters in voters' full ranked preferences – beyond their first choices – to fully understand their voting pattern in RCV elections. Third, we quantify the overall degree of racial polarization in elections with multiple racial groups by developing a novel statistic for racially polarized voting. Using precinct returns and cast vote records from mayoral elections in San Francisco and Oakland between 1994 and 2018, we find that the degree of racially polarized voting declined after RCV was implemented in both jurisdictions. Our results imply that implementation of RCV not only affects candidates' electoral strategies, but also how voters' preferences are expressed in elections.

Keywords: *Ranked-choice voting, racially polarized voting, cast vote records, electoral reform, rank data*

*This manuscript was prepared for the 2020 Southern Political Science Association Annual Meeting in Puerto Rico, January 9-11, 2020. The earlier version of this manuscript was presented at the 2019 Annual Midwest Political Science Association Conference. We would like to thank Chris Hughes, Shiro Kuriwaki, Randy Stevenson, and the participants of the two-days panel for ranked-choice voting for helpful comments. We also thank Erica Frazier for her assistance in data collection.

†Graduate student (corresponding author), Department of Political Science, Rice University. 6100 Main St, Houston, TX 77005 atsusaka@rice.edu.

‡Graduate student, Department of Government, Geogetown University. 3700 O St NW, Washington, DC 20057 tal76@georgetown.edu.

1 Introduction

In racially and ethnically divided societies, voting patterns have been often strongly determined by racial and ethnic divisions.[1] There is strong evidence of so-called racially and ethnically polarized voting from various contexts throughout the world (Bratton and Kimenyi, 2008; Bratton, Bhavnani and Chen, 2012; Shamir and Arian, 1982; Collet, 2005) and many studies have examined the effect of political institutions on the escalation and moderation of ethnic conflicts in politics (Huber, 2012; Brown, 2005; Issacharoff, 1992; Horowitz, 2000; Sisk, 1996; Reilly, 2001).

Over decades, political scientists have examined if ranked-choice voting (RCV), where voters can express their ranked preferences on multiple candidates, mitigates the degree of interethnic conflict in electoral competition (Horowitz, 2004; Fraenkel and Grofman, 2004; McDaniel, 2018; Coakley and Fraenkel, 2009; Coakley, 2009; Mitchell, 2014; Fraenkel, 2015). This expectation, called *moderation hypothesis*, states that RCV creates incentives for candidates to take more moderate policy positions and rely less on negative campaigning, which in turn promotes more crossover voting among voters of different racial and ethnic groups (Reilly, 2018). As a growing number of jurisdictions have recently adopted RCV in the U.S., it is of great interest to analyze the effect of implementing RCV on the degree of racially polarized voting in American elections. Despite the importance of the policy implication, previous research demonstrates remarkably mixed results on the moderation effect of RCV and, indeed, "there is little scholarly consensus" on the advocated moderation effect of RCV elections (Reilly, 2018, 207).

In this article, we provide the most direct test of the moderation effect of implementing RCV on the amount of racially polarized voting. In so doing, we contribute to the literature by employing several novel empirical approaches. First, we collect and analyze cast vote records where we have the known truth on individual rankings on available candidates in RCV elections. Second, drawing from the rank data literature, we explore potential clusters in voters' full ranked preferences – beyond their first choices – to fully understand their voting pattern in RCV elections. Third, we quantify the overall degree of racial polarization in elections with multiple racial groups by developing a novel statistic for racially polarized voting.

Specifically, we collect and analyze precinct returns and cast vote records from mayoral elections in San Francisco and Oakland between 1994 and 2018. This time period covers elections both before and after the

[1] In this article, we interchangeably use the term race and ethnicity, although they may have different meanings within some countries including the U.S. We also understand the difficulty in translating the concept of ethnicity in some countries to the one in other countries and treat them equivalently (e.g., treating the black and white racial difference in the U.S. equivalently as a difference in ethnicity in Malaysia). For this reason, future extensions must consider this point more deeply.

implementation of RCV in the two jurisdictions. Applying our newly proposed approach to RCV data, we find that the degree of racially polarized voting declined after RCV was implemented in both jurisdictions. Our results imply that implementation of RCV not only affects candidates' electoral strategies as recent studies discovered (Donovan, Tolbert and Gracey, 2016), but also how voters' preferences are expressed in elections.

This article proceeds as follows. Section 2 motivates our study by introducing previous research on RCV and racially polarized voting. Section 3 then discusses the theoretical expectation on the moderation effect of RCV on racial polarization. Next, Section 4 provides an overview of our research design and data sources. Section 5 details our approach to quantify the degree of racially polarize voting both in non-RCV and RCV elections. In Section 6, we present our findings on cast vote records analysis and the moderation effect of implementing RCV on racial polarization. Finally, Section 7 considers several implications of our study as well as future directions.

2 Ranked-Choice Voting and Ethnic Moderation

Ranked-choice voting (RCV) is an electoral system where voters can express their preferences on multiple candidates by ranking all or a subset of candidates, rather than only choosing the most desirable individual. While RCV can be implemented either with single-winner or multi-winner elections, we limit our scope of study to RCV elections with single-seat elections and call the latter combination of systems as the single-transferable vote (STV). Recently, a number of jurisdictions have adopted and implemented RCV elections throughout the U.S. as an alternative to the single-member district (SMD) with the plurality rule (Kimball and Anthony, 2018; Maloy, 2019).[2]

Under RCV, the winner of an election will be determined by a series of vote transfer given voters' ranked preferences on multiple candidates. When there is a candidate who obtains a majority support from voters in their first choices, the candidate becomes the winner and no vote transfer occurs. If such candidate does not exist in the first round, the candidate with the fewest votes will be eliminated and voters' ballots that went to the eliminated candidate will be transferred to their next preferred candidates in the second round. The sequence of vote transfer continues until there exists a candidate with the majority support. It is argued that this system gives voters more meaningful choices and *tends to* produce Condorcet winners, where a winner

[2]RCV has been introduced in over thirteen municipalities across seven states for local elections and the state of Maine for state and federal primary elections. Moreover, at least five subnational governments will implement RCV in upcoming elections.

is always preferred to another candidate in all possible pairwise comparison.

To illustrate the electoral system, we plot the number of rounds against the number of candidates in RCV elections in Figure 1. Note that under conventional SMD elections, the number of rounds is always one or two (if run-off exists). In contrast, the figure shows that over 47% of RCV elections experience more than three rounds. It seems also that the numbers of rounds and candidates are highly and positively correlated.

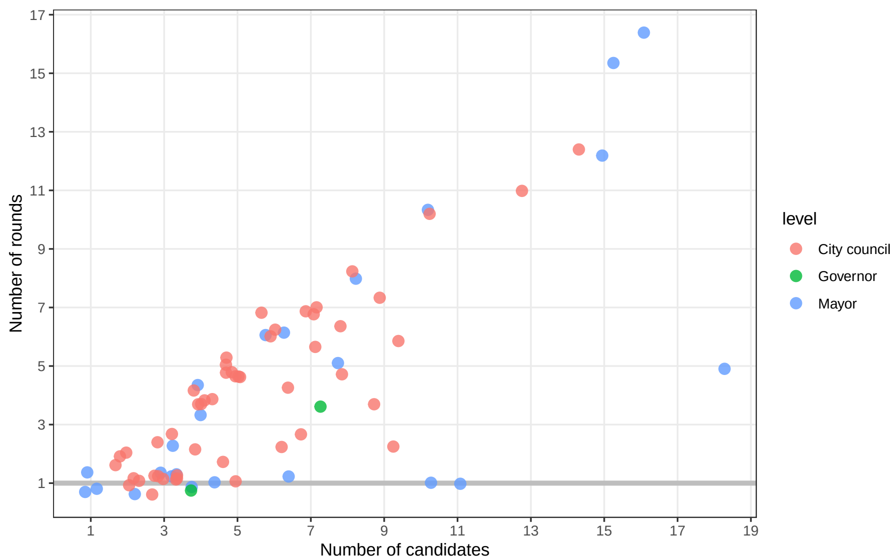

Figure 1: **Number of Candidates and Rounds in RCV Elections in the U.S.**
Note: This graph plots the number of rounds against the number of candidates in RCV elections in the U.S. The dots in the plot are jittered. The number of candidates does not account for write-in candidates. Data are collected from mayoral elections in San Francisco, Oakland, San Leandro, Berkeley (CA), Minneapolis, Saint Paul(MN), Portland (ME), Takoma Park (MD), Telluride (CO), gubernatorial elections in Maine, and city council elections in San Francisco and Oakland.

Importantly, the unique feature of RCV produces a wide range of expectations about the effects of RCV implementation on both elite and mass political behavior (Kimball and Anthony, 2018). For example, scholars have hypothesized and analyzed whether implementing RCV decreases the amount of negative campaigning (Donovan, Tolbert and Gracey, 2016; John and Douglas, 2017), boosts voter turnout (Schultz and Rendahl, 2010; McDaniel, 2016; Kimball and Anthony, 2016), and increases minority descriptive representation (John, Smith and Zack, 2018). In Figure 2, we present a logical chain of the expected effects of RCV implementation.

Here, we label the effects on candidate attributes and campaign style as "upstream" effects and the effects on election results and policy outcomes as "downstream" effects, leaving the effects on voting behavior in

RCV implementation
⇒ Candidates Attributes ⇒ Campaign Style ⇒ Turnout ⇒ Vote Choice ⇒ Winners ⇒ Policy Outcomes
Upstream *Downstream*

Figure 2: **A Logical Chain of the Effects of RCV Implementation**
Note: The above represents a logical sequence of the effects of RCV implementation on elite and mass political behavior.

the middle of the logical chain. When evaluating the causal effects of RCV implementation, it is crucial to recognize that any downstream effect cannot be assessed on its own and rather it must be considered along with a set of upstream effects in the logical chain. Otherwise, researchers must define an appropriate direct or indirect effect as an quantity of interest in causal inferences. For example, RCV implementation might affect voter turnout by affecting the civility of campaigns and simultaneously influencing voters' efficacy in elections.

Among these expectations, one of the most controversial and long-debated expectations is about the effect of RCV implementation on vote choice. Specifically, a large body of research has examined whether implementing RCV elections mitigates the degree of racial and ethnic polarized voting (Coakley and Fraenkel, 2017; Reilly, 2018; McDaniel, 2018). In this article, we call this expectation as the *moderation hypothesis* by borrowing from and modifying Reilly (2018, 209)'s "moderation thesis."

While numerous studies have explored the role of political institutions in the moderation of ethnic rivalry in ethnically divided communities (Huber, 2012; Brown, 2005; Issacharoff, 1992), perhaps the most heated debate in the literature is the one between Horowitz (1991*a*,*b*,*c*, 2004, 2006, 2007) and Fraenkel and Grofman (2004, 2006*a*,*b*, 2007). Horowitz argues that RCV produces incentives both for parties and the electorate to endorse ideologically more moderate candidates, which promotes cross-ethnic voting as well as interethnic accommodation and moderation. In contrast, Fraenkel and Grofman cast doubt on this proposition based on their formal model and draw evidence from Fiji elections, where native Fiji voters and those of Indian decent compete, to demonstrate that moderation does not necessarily happen. Other scholars have also scrutinized a similar effect of the electoral system on ethnic conflict in Northern Ireland (Coakley and Fraenkel, 2009; Coakley, 2009; Mitchell, 2014) and Africa (Fraenkel, 2015), while leading to mixed conclusions.

More recently, this important question has been revisited and examined in the context of American elections (Reilly, 2018; Robb, 2011; McDaniel, 2018). McDaniel (2018) provides some evidence that RCV does not alleviate the degree of racially polarized voting and may even increase interracial disparities in vote choice in American elections. Based on ecological inference and the difference-in-differences design, he demonstrates that the implementation of RCV in Oakland and San Francisco, California, does not reduce racial polarization in general and rather contributed to higher level of polarization between white and Asian electorates. On the contrary, drawing from Robb (2011), Reilly (2018) discusses positive evidence for the moderation hypothesis in Bay Area elections. Importantly, these substantially conflicting findings from the literature suggest that the moderation hypothesis has not been either accepted or rejected yet and further systematic investigation must be conduced.

Whereas the mixed conclusions might stem from underlying electoral environments,[3] a part of them could be derived from differences in empirical approaches. While McDaniel's pioneer work on RCV and racial polarization in American politics has made an important contribution to the literature, there are several limitations in his analyses. First, the author's analysis only incorporates voters' first choices and does not account for the full ranked preferences.[4] Next, his measurement of racially polarized voting only focuses on the differences in vote shares for the winners between a pair of racial groups (Hajnal, 2009; Hajnal and Trounstine, 2014) and does not fully evaluate the total degree of racial polarization among all racial groups. Consequently, his approach does not clarify whether voters' preferences are racially polarized in their ranked preferences as a whole. In fact, these concerns are not peculiar to McDaniel (2018) and rather they are common challenges in previous research on RCV and ethnic voting.

In this article, we contribute to the literature by testing the moderation hypothesis with several novel empirical approaches, in which we overcome the above challenges. First, we create a unique dataset from cast vote records where we know the ground truth about people's vote choice so that we can examine racial polarization in the entire ranked preferences. Second, drawing from the rank data literature, we estimate clusters in ranked ballots to understand which ranked preferences are close to each other. Third, we develop a new measurement of racially polarized voting in multiracial elections to make inferences about the total degree of racial polarization. Before discussing our analysis, we briefly elaborate a theoretical mechanism behind the moderation hypothesis in the next section.

[3]For example, Reilly suggests, "This limited spectrum of cases (most of them small and relatively obscure) and diversity of outcomes has meant that there is little scholarly consensus on core centripetal claims, with most scholars viewing the utility of preferential voting as highly contingent on facilitating social and demographic conditions" (Reilly, 2018, 207).

[4]We would like to thank Jason McDaniel for clarifying this important point and also for generously sharing his Bay Area Election Dataset.

3 The Moderation Hypothesis

As discussed above, in RCV elections, the winner is determined by an interactive allocation of ballots among ranked candidates. In each round, the candidate with the fewest votes will be eliminated and the ballots of voters who chose the eliminated candidate will be allocated to their next preferred candidates. The iteration continues until there exists a candidate who possess the majority of votes in that round (Burnett and Kogan, 2015; Grofman and Feld, 2004). Given this nature, Horowitz (1991*a*,*b*,*c*, 2004, 2006, 2007) argues that RCV tends to elect ideologically more moderate parties or candidates who can reach out to voters of different races and ethnicities for their second or lower ballots. Indeed, this causal mechanism is also assumed in the study of the effect of RCV implementation on the degree of negative campaign (Donovan, Tolbert and Gracey, 2016; John and Douglas, 2017; Kimball and Anthony, 2018).

While this mechanism is often discussed without mention on underlying assumptions, we argue that it is crucial to understand a set of assumptions which are basis of the moderation hypothesis (Fraenkel and Grofman, 2004, 2006*a*,*b*, 2007). First, it is assumed that voters always support their co-ethnic parties or candidates in the first choice (**Co-ethnic Voting**). Second, those parties and candidates are assumed to be able to win elections if and only if they can collect additional ballots from voters of different groups in the second or lower choices (**No Dominant Group**). Third, moderate parties and candidates are more apt to gain lower ballots from people of different ethnicities given voters' strategic consideration and preferences for moderate politicians (**Preferences for Moderate Parties**).

It must be emphasized that the moderation hypothesis is only plausible when the above assumptions are satisfied. Hence, even when evidence against the moderation hypothesis is provided, it does not immediately mean that the moderation hypothesis itself is illogical or incorrect. Rather, it would be more prudent to consider the plausibility of the assumptions before drawing a conclusion about the hypothesis. Indeed, after finding a null result for the moderation hypothesis, Fraenkel and Grofman raise a question about the last assumption and contend that moderation does not occur when moderate parties are not already popular in a given society.[5]

In this article, we contend that the above three assumptions are more or less plausible in elections examined below. Assuming that the three assumptions hold, we draw the following hypothesis.

[5] It must be emphasized that previous discussion assumes the presence of only two major ethnic groups in a given society. Thus, the future extension of this article must include a formal analysis of multiracial societies building on Fraenkel and Grofman (2004).

H$_1$: *Switching from the plurality rule to ranked-choice voting* ***reduces*** *the degree of racially polarized voting.*

For completeness, our null hypothesis is as follows:

H$_0$: *Switching from the plurality rule to ranked-choice voting* ***does not affect*** *the degree of racially polarized voting.*

The next two sections illustrate our empirical strategy to test the moderation hypothesis.

4 Research Design and Data

4.1 Population and Quantities of Interest

To test whether the implementation of RCV reduces the degree of racial polarization, we define our population of interest as a set of *electoral contests* which could have been conduced under non-RCV (or SMD) and RCV elections. Given our population of interest, we now formally define our quantities of interest.

Let Y_i be a continuous random variable for the outcome, denoting the degree of racially polarized voting at contest i. Let D_i denote a binary random variable for the treatment status, representing if the same contest was an RCV election. Based on the potential outcomes framework, let us define $Y_i^{d=1}$ as a potential outcome for the contest had it been assigned to a jurisdiction with RCV, and $Y_i^{d=0}$ as a potential outcome for the same contest had it been assigned to a jurisdiction without RCV. Here, we assume that the consistency assumption holds as $Y_i = D_i Y_i^{d=1} + (1 - D_i) Y_i^{d=0}$, where the observed outcome corresponds to a potential outcome under the observed treatment status. Assuming an additive effect measure, we define the average treatment effect on the treated (ATT) as:

$$\tau_{ATT} = \mathbb{E}[Y_i^{d=1} - Y_i^{d=0} | D_i = 1]. \tag{1}$$

In practice, since we use observed outcomes of contests, we are concerned with the ATT conditional on a set of covariates.

$$\tau_{ATT}^{X} = \mathbb{E}_X \left[\mathbb{E}[Y_i^{d=1} - Y_i^{d=0} | D_i = 1, \mathbf{X} = \mathbf{x}] \right], \tag{2}$$

where $\mathbf{X}$ is a vector of p covariates and $\mathbf{x}$ denotes a vector of specific covariate values, defined over the p-dimensional covariate space $\mathcal{X}$. We assume that the conditional ignorability holds such that within a multidimensional strata of background characteristics the potential outcomes of contests in the treatment condition and control condition and the treatment assignment are statistically independent. Moreover, our quantity of interest is a "contemporaneous" effect of the implementation of RCV on the degree of racial polarization. This means that we are primarily concerned with the causal effect of the treatment on the outcome evaluated at time t instead of other long-term effects caused by the RCV implementation.

4.2 Cast Vote Records in RCV Elections

Given our quantities of interest (ATT), an ideal research design would be to perform a block randomized controlled trial in which the two electoral systems are randomly assigned to voters of different racial groups in actual elections. Unfortunately, such field experiment cannot be administered due to practical and ethical constrains and we instead rely on observational data. Specifically, we perform an observational study using a time-series cross-sectional (TSCS) data where the variation in electoral systems in actual elections is exploited to estimate the causal effect of interest.

In this preliminary analysis, we use cast vote records and precinct returns from San Francisco and Oakland mayoral elections to derive our quantities of interest. Cast vote records, also known as ballot image logs, are a digitally recorded list of numbers, containing indicators for elections, voters, precincts, vote choices among other information. Cast vote records are emerging type of data in which we know exactly how voters chose – and in our case ranked – among multiple candidates (Gerber and Lewis, 2004; Herron and Lewis, 2007; Park, Hanmer and Biggers, 2014; Alvarez, Hall and Levin, 2018; Kuriwaki, 2019).

San Francisco and Oakland are two of the first jurisdictions that have implemented RCV elections in the U.S.. For pre-intervention periods, precinct returns are examined in San Francisco (1995, 1999, 2003) and in Oakland (1998, 2002, and 2006), respectively. For post-intervention periods, cast vote records are scrutinized in San Francisco (2011, 2015, and 2018) and Oakland (2010, 2014, and 2018). The two jurisdictions are also analyzed in McDaniel (2018), and we incorporate a longer time period for post-intervention contests in our study.

4.3 Treatment Variable

Our treatment variable is a binary indicator for elections conducted under the RCV rule. Figure 3 portrays the distribution of the treatment variable between 1994 and 2020. We chose this time period in order to secure enough units before and after the implementation of RCV which has occurred between 2007 and 2010. In Figure 3, the dark color represents treated units (i.e., RCV elections) and the light color depicts control units (i.e., non-RCV elections). Our data do not contain any reversal in treatment status where RCV jurisdictions returned to non-RCV jurisdictions. It should be noted that while the outcome variable is measured at the unit level, the treatment assignment occurs only at the municipality level. Across the treatment status, we do not observe any clear pattern in the number of candidates, number of rounds, and incumbency status. The details of fourteen elections are discussed in Appendix A.

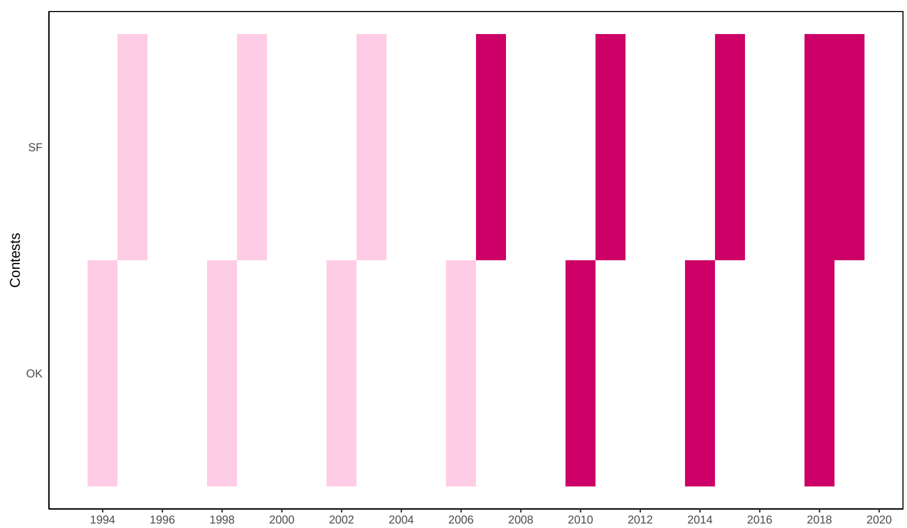

Figure 3: **Distribution of the Treatment Variable**
Note: This figure visualizes the empirical distribution of the treatment variable. The dark pink represents RCV elections (treated), whereas the light pink illustrates non-RCV elections (control).

4.4 Outcome Variable

Our outcome variable is a continuous measure for the overall level of racially polarized voting in RCV and non-RCV elections with voters of multiple racial groups. In this article, we consider whites, blacks, asians, and hispanics as primary categories for the raical groups of interest. To construct the outcome measure for RCV elections, we first estimate clusters in cast vote records to reduce the number of possible ways that voters can cast ballots into three clusters. We next estimate the proportion of voters from each group belonging to each cluster based on ecological inference. Finally, we map the multidimensional metrics (i.e., multiple groups and multiple clusters) onto a uni-dimensional statistic that summarizes the overall level of racial polarization in an election. For non-RCV elections, we consider three clusters, voting for the winner, runner-up, and other candidates, and apply the same measurement strategy to construct the outcome variable. Since RCV data are relatively new and highly complex data, the details in how we measure the outcome variable are elaborated in the next section.

4.5 Potential Confounders

For valid causal inferences, it is crucial to carefully consider the treatment assignment mechanism and potential confounders which affect both the treatment status and outcome. In our study, it implies that we must understand how and where RCV elections are implemented and how the jurisdictions' selection into the treatment group is related to the degree of racial polarization. In other words, we must perform our analysis conditional upon a set of all confounders which affect *both* the RCV implementation and racial polarization. *While we do not condition on these confounders in our analysis below, let us describe how we proceed in our future extensions here.*

In this study, we consider three variables for confounding factors, which include the proportion of non-Anglo eligible voters, the level of residential segregation, and the overall turnout rate. Figure 4 depicts a possible causal mechanism that we consider to be plausible.

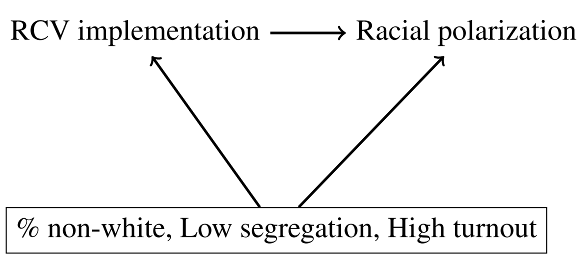

Figure 4: **Directed Acyclic Graph for the Effect of RCV Implementation**
Note: This figure presents a directed acyclic graph (DAG) for the causal relationship of interest.

The selection of the three variables is based on our qualitative research on the adoption mechanism of RCV elections in the Bay Area. In particular, we observe that RCV adoption has typically been driven by feelings of under representation both among specific racial groups (%nonwhite) and the general public (Low

Turnout). In the US, these concerns have typically been resolved via the creation of single member districts and districting regimes that allows each voting block some degree of representation. District based solutions also have the attractive feature of eliminating turnout concerns, since a geographically bounded population of interest will be able to elect their representative no matter how low turnout in their district is, although districts do not solve the problem of chronically low turnout and may exacerbate it (Brace et al., 1995).

While the single member district revolution at the State and Federal level was largely complete by the mid-20th century (Cain, Donald and McDonald, 2005), local elections in the U.S. have often been overlooked and many still use election systems such as at large block voting, which drastically under represent minority communities even when such communities are fairly large and well organized (Taebel, 1978). The Bay Area cities that adopted RCV between 2000 and 2010 did so largely because earlier attempts to reform such systems, via single member districts (SMDs), runoffs, top two primaries, or at-large numbered posts (a variant of block voting where each seat is treated as its own election, but all are elected from the same electorate), were insufficient.

Thus the importance of our confounding variables can be explained by the confluence of circumstances in which they cause single winner districts to fail as a remedy for achieving representation of all key stakeholders. District based remedies are dependent on a degree of segregation between relevant stakeholder groups to function (Vedlitz and Johnson, 1982), and the Bay Area has been the epicenter of a growing trend in which segregation between certain groups (such as Asian and White) has broken down, and where areas are increasingly defined not by a majority and a minority group, but by a set of 3-5 ethnic groups where the plurality ethnic bloc may contain only 30-40% of the electorate (Frey, 2018).

Runoffs are a direct attempt to solve this problem in that they force groups to form coalitions via a narrowed choice set, but runoffs also lower turnout in ways that can disproportionately effect minority groups and lead to a white plurality group candidate winning over a multi-racial coalition candidate. RCV emerges in these contexts as an attempt to gain the beneficial effects of runoffs on group coordination without the problem of lowered turnout caused by multiple elections.

In other words, we expect and have seen RCV emerge in contexts where turbulent dynamics around race and election results are already present. Importantly, the three factors are also expected to influence the degree of racial polarization in elections. However, this does not imply perfect correlation between RCV implementation and racially polarized outcomes. A number of cities in the Bay Area and around California facing California Voting Rights Act litigation have considered utilizing RCV and ultimately rejected it, and there was likely a bandwagon effect among some cities that were part of the initial wave of implementation. Furthermore, in the decade since RCV was implemented in Alameda county and more than 15 years since RCV was implemented in San Francisco, political cleavages have changed significantly, meaning we should still observe significant variation in racial dynamics within RCV cities. We believe including controls for these variables is sufficient to resolve these confounding issues.

5 Measuring Racial Polarization in RCV Elections

In this section, we detail our approach to quantify the degree of racially polarized voting in RCV and non-RCV elections. Here, we point out that to fully understand the complex structure of RCV data it is necessary to draw ideas and tools from the rank data literature (Fligner and Verducci, 1993; Marden, 1995; Alvo and Yu, 2014; Liu et al., 2018). Despite the fact that RCV elections are perfect applications for rank data analysis, almost no political science research has adopted the rank data approach to RCV data. Hence, we contribute to the literature by bridging the gap between the RCV scholarship and the statistics field to study voter preferences under RCV and non-RCV elections. For a more general introduction to the field, see Atsusaka (2019). Below, we focus on the concept of clustering in rank data and introduce our empirical strategy for clustering our cast vote records.

5.1 Clustering Individual Rank Data

To quantify racial polarization in ranked preferences, it is necessary to understand which rankings are similar and dissimilar to each other. Especially, clustering multiple rankings to classify voting patterns is essential as the total number of rankings becomes as large as $N!$ (in compulsory ranking systems) or $2\binom{N}{C}C! + \sum_{n=1}^{C-1}\binom{N}{C-n-1}(C-n-1)!$ (in optional ranking systems), where N is the total number of candidates and C refers to the total number of rankings that voters can express with $C \leq N$. In the latter, voters can engage in so-called "plumping" by only raking a single candidate or subset of candidates to represent more nuanced ranked preferences (Coakley and Fraenkel, 2017, 676).[6]

To illustrate why this is problematic, let us first describe how researchers usually quantify racial polarization in non-RCV elections. Table 1 represents a contingency table for a hypothetical election where there

[6]For example, when N=5 and C=3, plumping becomes a ranked preference such as (A, □, □) and (A, C, □). Such preferences are clearly different from (A, C, E) as the voter does not want F as other unranked candidates. In other words, plumping is not a partial ranking, where lower ranks are randomly missing, but a full ranking. In Section 6, we show that about 10 to 55% of voters indeed engage in plumping in our data.

are two groups (Asian and White) and two candidates (A and B). There are 100 voters in each group and Candidates A and B received 105 and 95 votes, respectively. These information are expressed by the numbers on row and column margins. Now, if we look at the number in each cell, we notice that 90% of asian voters preferred Candidate A, whereas only 15% of white voters preferred the same candidate. Similarly, Candidate B was preferred by 85% of white voters and only 10% of asian voters. Given that, researchers usually infer that vote choice was racially polarized in the election. While we employed the simplest example with two groups and two candidates, similar inferences can be made for multiple groups and multiple candidates.

	Candidate A	Candidate B	
Asian	90	10	100
White	15	85	100
	105	95	

Table 1: **An Example of Racial Polarization in non-RCV Elections**
Note: This table represents a typical contingency table used in the analysis of racially polarized voting in non-RCV elections with two candidates and two groups of interest. Row margins are the total number of voters in the two groups, whereas column margins show the total number of votes that the two candidates received.

Now, let us describe what would happen if we attempt to draw similar inferences for RCV elections. Suppose that we have five candidates (N=5) (A,B,C,D,E) and voters can cast ballots up to three ranks (C=3) in a hypothetical RCV election. Voters are now asked to express only one ranked preference (e.g., (A, C, E), (B, D, A), etc) among all the possible permutations of candidates. If the optional ranking system is employed, the total number of ways voters can vote becomes $2\binom{5}{3}3! + \sum_{n=1}^{2}\binom{5}{2-n}(2-n)!$=126. Assuming that there are two groups (Asian and White) as before, the election result is now represented by the following **2** $\times$ **126** table.

	ABC	ABD	ABE	$\cdots$	E□□	
Asian	1	1	0	$\cdots$	1	100
White	1	0	0	$\cdots$	4	100
	2	1	0	$\cdots$	5	

Table 2: **An Example of Racial Polarization in RCV Elections**
Note: This table represents a contingency table that can be potentially analyzed to study racially polarized voting in RCV elections. On columns, all possible rankings that voters can potentially express are listed.

One may rapidly notice that there are at least two critical limitations in this naive approach. First, since we now have 126 unique "choices" and each cell contains a small number of voters, it is extremely difficult to draw an intuitive inference about how asian and white voters behave differently. Second, this approach treats (A,B,C) vs. (A,B,D) and (A,B,D) vs. (E,□,□) as *equally different* ranked preferences, despite the fact that the first two choices both include A and B in the same ranks and the last choice does not include any of A, B, and D. In order to draw substantively meaningful inferences about racial polarization, it is then necessary to somehow *reduce* the number columns or *cluster* multiple ranked preferences into few meaningful categories. We must note that this problem occurs regardless of data sources (i.e., whether we use individual survey data or aggregate data).

To overcome this *column reduction problem*, we propose two different approaches: agnostic column reduction and substantive column reduction. The agnostic column reduction relies on clustering algorithms in rank data analysis to categorize multiple ranked preferences based on how people actually voted. For example, a given data might have many voters who chose (A,B,C) and (A,C,B), while it also has many voters who chose (D,E,F) and (D,F,E). Then, the agnostic column reduction estimates two major clusters in the data, namely, ABC vs. DEF. While a recent development in rank data analysis offers a way to estimate potential clusters in ranked data, it does not distinguish plumping (e.g., E,□,□) from randomly missing values (e.g., E,*NA*,*NA*) (Vitelli et al., 2018). To overcome this challenge and draw more valid inferences from RCV data, we develop a method to estimate clusters in RCV data with plumping in our future work.

The substantive column reduction employs information on candidate attributes and clusters multiple ranked preferences into substantively same categories. For example, suppose we know that A and B are leftist candidates, C and D are moderates, and E is a rightist candidate. We can then consider (A,B,C), (B,A,C), (A,B,D), and (B,A,D) as the substantively same preference of (Left, Left, Moderate). By coding candidates' policy position, race and ethnicity, and partisanship, we estimate racial polarization based on different candidate attributes.

In this article, we only apply the agnostic column reduction approach. The product of such column reduction is represented by Table 3. In future research, we compare estimates based on both the agnostic and substantive column reductions and examine whether a substantive conclusion about the moderation hypothesis differs based on different approaches.

	Cluster 1	Cluster 2	Cluster 3	
Asian	80	10	10	100
White	5	85	10	100
	85	95	20	

Table 3: **An Example of Clustering Multiple Rankings**
Note: This table represents a contingency table that can be potentially analyzed to study racially polarized voting in RCV elections. On columns, three clusters of similar rankings are listed.

5.2 Estimating Clusters in Cast Vote Records

Having established the need for data reduction in cast vote records, we now illustrate how we perform the agnostic column reduction. To date, very few studies have considered clusters in rank orderings in RCV elections (Gormley and Murphy, 2008; Gormley, Murphy et al., 2008). To facilitate the discussion, let us first introduce several notations and concepts for rank data analysis.

Rank data arises whenever a number of units put rankings on multiple elements.[7] In the rank data literature, such units are called *judges* or *assessors* and these elements are referred to as *items*. Judges can be individuals, voters, survey respondents, leaders, groups, organizations, or countries. Items can be candidates, politically salient groups, options, public figures, political parties, or types of representation. Judges' preferences are then represented by how they "permute" available items. In our case, judges are voters and items are candidates. In RCV elections, voters express their preferences by choosing multiple candidates in a specific order.

A *ranking* is thus considered as a projection from items to permuted numbers. Let $\mathcal{A} \in \{A_1, ..., A_n\}$ be a finite set of items (e.g., candidates) on which a judge (e.g., voter) puts ranking according to some criteria. Let $R_i \in \{1, ..., n\}$ be a rank that the judge assigns to each item A_i. Then, a full ranking of all items $\mathbf{R} = (R_1, ..., R_n)$ becomes a mapping: $\mathcal{A} \rightarrow \mathcal{P}_n$, where $\mathcal{P}_n$ is the space of n-dimensional permutations. The assessor is said to prefer A_i to A_j when $R_i < R_j$ and the most preferred item is represented as A_1 (i.e., A_1 has the *highest* rank).

Next, let $\mathbf{R}_j$ be the full ranking given by judge j, where $j \in \{1, ..., N\}$. Then, rank data can be considered as a matrix where judges are listed on rows and items are located on columns. Given full rankings provided by multiple judges, one of the goals of rank data analysis is to estimate so-called the shared true consensus ranking, which is a ranking (i.e., list of numbers) that most judges agree upon. Specifically, various probabilistic models have been considered in order to estimate the consensus ranking.

[7]Rank data are also referred to as ranking data or ranked data.

While there exist four types of probabilistic models in the rank data literature (Alvo and Yu, 2014; Liu et al., 2018), we consider a class of distance-base models called the Mallows model (Mallows, 1957; Diaconis, 1988). The Mallows model assumes that the probability of observed ranking decays as the "distance" between the assessor's ranking and what is called the shared true consensus increases. For the Mallows model, the likelihood of function becomes,

$$P(\mathbf{R}_1, ..., \mathbf{R}_N | \alpha, \boldsymbol{\rho}) = \prod_{j=1}^{N} \frac{1}{Z_n(\alpha)} \exp\left\{ -\frac{\alpha}{n} \sum_{j=1}^{N} d(\mathbf{R}_j, \boldsymbol{\rho}) \right\}, \tag{3}$$

where $\boldsymbol{\rho} \in \mathcal{P}_n$ is a vector of location parameters which represent the shared true consensus ranking among assessors, $\alpha > 0$ is a scale parameters, $Z_n(\alpha)$ is a partition function (i.e., normalizing constant), and $d(\mathbf{R}_j, \boldsymbol{\rho})$ is a distance between an observed ranking by judge j and the shared consensus ranking. While various distance functions have been proposed in the literature, we use the footrule distance in this article.

Given this likelihood function, our goal is to make inferences about the true consensus ranking among assessors $\boldsymbol{\rho}$. Recently, Vitelli et al. (2018) proposed the Bayesian Mallows model (BMM) to compute the posterior distribution of $\boldsymbol{\rho}$ given the observed likelihood and specified prior distributions such that:

$$P(\alpha, \boldsymbol{\rho} | \mathbf{R}_1, ..., \mathbf{R}_N) \propto \prod_{j=1}^{N} \frac{1}{Z_n(\alpha)} \exp\left\{ -\frac{\alpha}{n} \sum_{j=1}^{N} d(\mathbf{R}_j, \boldsymbol{\rho}) \right\} P(\alpha) P(\boldsymbol{\rho}) \tag{4}$$

Here, a uniform prior distribution is applied to the location parameter $\boldsymbol{\rho}$, $P(\boldsymbol{\rho}) = n!^{-1}$ and an exponential prior is chosen for the scale parameter α, $P(\alpha) = \lambda_\alpha \exp(\lambda_\alpha) \mathbf{1}_{[0,\infty)}(\alpha)$, where the hyperparameter λ_α is treated as a tuning parameter. With a choice of the distance function, the model is estimated via the Metropolis-Hastings algorithm presented in Vitelli et al. (2018).

Now recall that our goal here was to estimate possible clusters in cast vote records for RCV elections. To estimate clustering in RCV data, we employ a mixture of the BMM, where each assessor is assigned to one of C clusters by $z_1, ..., z_N \in \{1, ..., C\}$ and the consensus ranking within the same cluster is denoted by $\boldsymbol{\rho}_c$ and scale parameter by α_c. The likelihood for the mixture becomes,

$$P(\mathbf{R}_1, ..., \mathbf{R}_N | \{\boldsymbol{\alpha_c}, \boldsymbol{\rho_c}\}_{c=1,...,C}, z_1, ..., z_N) = \prod_{j=1}^{N} \frac{1}{Z_n(\alpha_{z_j})} \exp\left\{ -\frac{\alpha_{z_j}}{n} \sum_{j=1}^{N} d(\mathbf{R}_j, \boldsymbol{\rho}_{z_j}) \right\} \tag{5}$$

Inferences for the location parameters are made similarly by the MCMC algorithm with the choice of prior distributions. The optimal number of clusters is determined by the within-cluster sum-of-squares criterion for model selection (Vitelli et al., 2018, Section 6.3). However, for our analysis below, we *ex ante* determine that $C = 3$ to reduce the complexity of our outcome measure. In future research, we will tackle the problem of optimal number of clusters in the agnostic column reduction.

Finally, we should note that while the above clustering algorithm is assumed to be applied to the entire rank data, using all judge in clustering turns out to be computationally quite challenging. Therefore, in our study, we first performed a stratified sampling of size 20000 per contest, where precincts were used as strata. The sample size is approximately 10-15% of the entire data. We then applied the clustering algorithm to the sampled data and obtained estimated clusters for sampled voters. Finally, by using the matching between unique rankings and estimated clusters in the sample, we imputed estimated clusters for the entire dataset.

To demonstrate our approach, let us briefly describe our results from the 2018 San Francisco mayoral special election. In the special election, there were eight qualified candidates and voters were able to rank up to top three preferred candidates. In total, there were 456480 valid votes, among which there were 376 unique rankings out of 1018 possible rankings. Using the approach discussed above, we estimated three clusters from the data. The estimated results are represented in Table 4. In the special election, London Breed (in **bold**) won by 50.55% vote share over the runner-up Mark Leno whose final vote share was 49.45% after the ninth round of vote transferring. The third place and four place were Jane Kim and Angelra Alioto, respectively. Clusters 1 to 3 consist of 43%, 18%, and 39% of valid ballots, respectively. In our analysis, thus, we treat (Leno, Kim, Alioto), (Kim, Leno, Breed), and (Breed, Leno, Kim) as three different "vote choices" in which we examine the degree of racially polarized voting.

Mean Rank	Cluster 1	Cluster 2	Cluster 3
1	Mark Leno	Jane Kim	**London Breed**
2	Jane Kim	Mark Leno	Mark Leno
3	Angela Alioto	**London Breed**	Jane Kim
4	**London Breed**	Ellen Lee Zhou	Angela Alioto
5	Ellen Lee Zhou	Angela Alioto	Michelle Bravo
6	Richie Greenberg	Amy Farah Weiss	Ellen Lee Zhou
7	Amy Farah Weiss	Michelle Bravo	Amy Farah Weiss
8	Michelle Bravo	Richie Greenberg	Richie Greenberg

Table 4: **Estimated Clusters with Candidates' Mean Rank**
Note: This table shows the results of the clustering algorithm applied to 2018 San Francisco mayoral special election. London Breed, shown in bold, is the winner of the election. The horizontal line implies that voters were only allowed to rank up to three candidates in the election.

For non-RCV elections, we similarly consider three clusters, where the winner, runner-up, and rest of other candidates are classified as the first, second, and third clusters.

5.3 Summarizing Racial Polarization in Multiracial Elections

Political scientists and legal scholars have long struggled with the concept of racially polarized voting and its empirical measure (Elmendorf, Quinn and Abrajano, 2016; Ansolabehere, Persily and Stewart III, 2009, 2012; Tam, 1995; Collet, 2005; Grofman and Migalski, 1988; Engstrom, 2015; Bullock III and Campbell, 1984; Hasen, 2018; Grofman, 1991; Lupia and McCue, 1990; Cain and Zhang, 2016). This is especially true when researchers observe voters of multiple racial groups because most analyses are based on the assumption of biracial elections (Greiner, 2011). While it is still helpful to quantify the difference between a pair of racial groups (McDaniel, 2018), we believe that having summary statistics for the total degree of racial polarization across all racial group helps us evaluate the effect of RCV more effectively.[8]

In this article, we define racially polarized voting as a class of voting patterns where voters' race and vote choices are highly correlated.[9] Our proposed measure of racially polarized voting is a function of the absolute difference between the proportion of the racial group r belonging to cluster k and the proportion of the rest of the electorate belonging to the same cluster. We calculate such difference between one race and all other racial groups for all cluster and then take the average of such difference over all racial groups.

Formally, we define our racial polarization measure δ as follows:

$$\delta \equiv \sum_{k=1}^{K} w_k \left(\frac{1}{R} \sum_{r=1}^{R} \Big| \mathrm{P}_{r,k} - \mathrm{P}_{-r,k} \Big| \right) \tag{6}$$

where $\mathrm{P}_{r,k}$ represents the proportion of racial group $r \in R$ belonging to cluster $k \in K$, $\mathrm{P}_{-r,k}$ denotes the same proportion for the all other racial groups combined, and w_k is a weight for each cluster. For the weight, we use cluster vote shares both in non-RCV and RCV elections. When $w_k = 1/K$, the quantity becomes an unweighted measure. Below, we both analyze weighted and unweighted measure of racial polarization. The polarization measure δ takes 0 *either* if all racial groups agree on belong to the exact same cluster *or* if

[8] It must be emphasized that we do not claim that our measurement could be a relevant statistic in the plethora of redistricting and the Voting Rights Act cases. Our only focus here is to evaluate whether the proposed measure is an empirically valid statistic, and not a legally meaningful criterion.

[9] It is crucial to recognize that some scholars conceptualize such phenomena as racially correlated voting and separate them from racially polarized or racially caused voting where voters' true preferences, caused by race, are highly correlated with their vote choices (Greiner, 2011; Elmendorf, Quinn and Abrajano, 2016). We support this distinction both for empirical and legal studies; nevertheless, we use the conventional term of racially polarized voting in this article. Again, our definition does not include voters' preferences and their ties with their racial identification.

members of all racial groups belong to every cluster with the same proportion.

5.4 Ecological Inference

So far, we have assumed that we observe candidate vote shares or cluster proportions by racial groups $P_{r,k}$ and $P_{-r,k}$. In practice, however, secret ballots will not allow us to directly measure these quantities. In order to overcome the measurement problem, we perform ecological inference to estimate the target quantities from two known aggregated information of precinct-level racial makeup and cluster vote shares (King, 1997). Table 5 illustrates our strategy.

	Cluster 1	Cluster 2	Cluster 3	Non-Vote	
Asian	?	?	?	?	200
White	?	?	?	?	200
Hispanic	?	?	?	?	130
Black	?	?	?	?	90
	170	110	40	300	

Table 5: **An Example of Ecological Inference Data**
Note: This table represents a typical contingency table to which ecological inference is applied. Here, only row and column margins are actually observed, and table cells are not observed.

From cast vote records and precinct returns, we observe the total number of voters who support each candidate or belong to each cluster at the precinct level. These quantities are represented in the column margin. Using data from the U.S. Census and American Community Survey, we also obtain the distribution of voters by groups at the precinct level, which are shown in the row margin. For Oakland and pre-RCV San Francisco elections, we use precinct-level racial makeup created by McDaniel (2018). The goal of ecological inference is to estimate the quantities in the table cell based on the two marginal distributions of election results and the racial composition. Specifically, we calculate the municipality level quantities in the table cell from precinct-level estimates, which we then plug-in in Equation (6).

For estimation, we employ the Multinomial-Dirichlet ecological inference (MD) model proposed by Rosen et al. (2001) and implement it with `eiPack` (Lau, Moore and Kellermann, 2007) in R (R Core Team, 2019).[10] For valid ecological inference in the MD model, it is important to make sure that the cluster vote shares in one precinct is statistically independent from the cluster vote shares in other precincts. In order to satisfy this condition, we incorporate the proportions of white, black, asian, and hispanic voters as a set

[10]Future extensions must include estimates from the QG model that directly models the internal cell counts as opposed to the MD model that models the proportions instead (Greiner and Quinn, 2009, 2010).

of covariates in the MD model. Technical details are discussed in Appendix C. We then derive the mean estimate and 95% credible intervals of our quantities of interest based on 50000 draws from the posterior distributions via the MCMC including 10000 first draws as burn-in periods.

6 Empirical Findings

In this section, we report our main empirical results. First, we discuss our cast vote records and the results of our clustering algorithm applied to our RCV data (Section 6.1). Second, we provide an empirical test for the moderation hypothesis (Section 6.2).

6.1 Clusters in RCV Data

We begin by analyzing the cast vote records that we directly collected from San Francisco Department of Elections and Almeda County Registrar of Voters. Before analyzing data, we excluded both over votes and under votes, where voters cast more than one ballots to the same candidate and they do not cast ballot to any qualified (non-write-in) candidates, respectively.

In all RCV elections we analyzed, voters were allowed to rank up to three candidates, whereas the number of candidates is often two to four times larger than the legal restriction.[11] Hence, our cast vote records are *top-3 ranking data*. Moreover, a substantial portion of voters did not rank all of three candidates. Column (3) of Table 6 reports the percentage points of voters who rank all of the three positions. We find that about 10 to 55% of voters only selected one or two candidates to express their preferences. Columns (4)-(6) also provide descriptive statistics for the number of unique observed rankings, total number of possible rankings, and the ratio of the two statistics. Here, the total number of possible rankings includes the possibility of "plumping" where voters only choose a single or two candidates. The ratio statistic implies that among all possible ways of casting ranked ballots only 22 to 49% of rankings are actually observed.

We now apply the clustering algorithm based on the Bayesian Mallows model to our data. In all clustering algorithm for rank data, the number of clusters must be *ex ante* specified by researchers. Generally, it is suggested that analysts run such algorithm with different number of clusters and choose the number of clusters which minimize the within-cluster distances between the consensus ranking and observed rankings (Liu et al., 2018). However, we only consider three clusters in each data for simplifying our analysis and

[11] Since 2019, San Francisco mayoral elections allow voters to rank all candidates.

City	Year	%Full	#Unique Rankings	#All Rankings	Ratio	Cluster shares
Oakland	2018	56.3	673	1992	0.34	(0.52, 0.30, 0.18)
Oakland	2014	73.6	1475	8178	0.18	(0.57, 0.29, 0.14)
Oakland	2010	71.2	740	1992	0.37	(0.51, 0.46, 0.03)
San Francisco	2018	89.5	376	1018	0.37	(0.43, 0.39, 0.18)
San Francisco	2015	44.6	211	428	0.49	(0.53, 0.49, 0.03)
San Francisco	2011	72.7	1759	8178	0.22	(0.55, 0.32, 0.13)
San Francisco	2007	34.8	974	3446	0.28	(0.55, 0.32, 0.13)

Table 6: **Findings from RCV Cast Vote Records**
Note: This table summarizes cast vote records from RCV elections.

ensuring comparability with non-RCV elections, where we consider vote shares for the winner, runner-up, and the rest of other candidates combined. In our future extension, we take up on this issue of selecting the number of clusters in more detail. The shares of estimated clusters for each election are shown in Column (7) of Table 6. In all election, we discovered three clusters with varying sizes. These cluster shares are used as weights on racial polarization in the next section.

6.2 Evidence for the Moderation Hypothesis

To test the moderation hypothesis, an ideal identification strategy would be to rely on the recently proposed matching-based approach with Time-Series Cross-Sectional data (Imai, Kim and Wang, 2018). However, this approach requires a decent number of control units (i.e., non-RCV elections), on which we are still under process of data collection. Thus, in this article, we turn to a simple visual inspection and descriptive analysis of the effect of RCV implementation on the degree of racially polarized voting.

First, we plot the distribution of the racial polarization measure over time in Oakland and San Francisco in Figure 5. In the figure, dots represent weighted estimates and triangles visualize unweighted estimates, where we treat all cluster equally regardless of vote share. Pre-RCV elections are shown in gray and RCV elections are shown in purple. 95% credible intervals are also reported as error bars.

For weighted estimates, RCV elections seem to have lower level of racially polarized voting, except for 1994 and 1998 elections in San Francisco, which have approximately the same amount of polarization as RCV elections. For unweighted estimates, RCV elections seem to have lower level of racial polarization in San Francisco, but not in Oakland. These results imply that the RCV implementation seems to be negatively related to the degree of racially polarized voting, but further investigation is required to make a causal claim.

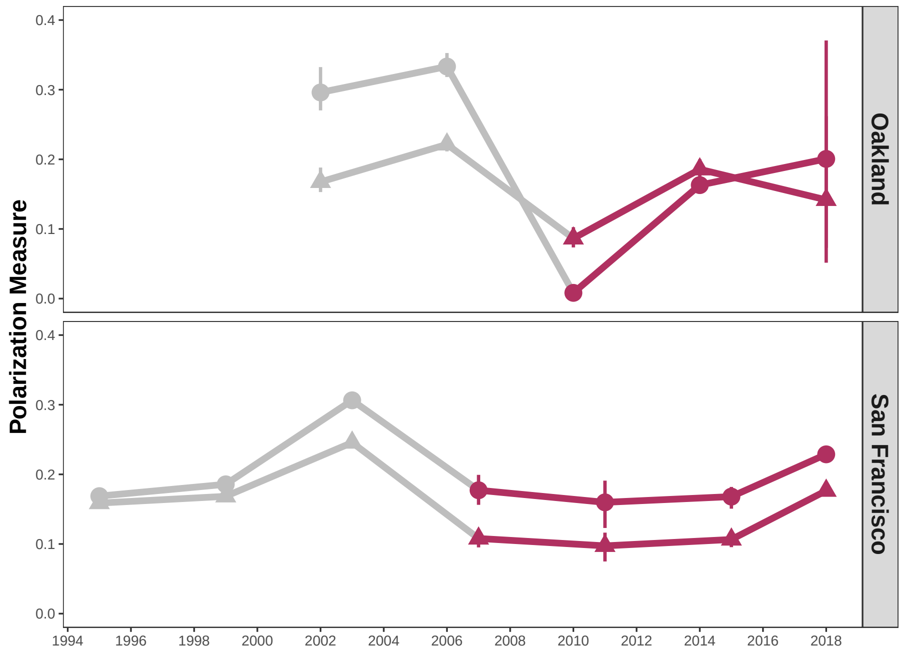

Figure 5: **Racially Polarized Voting Before and After the RCV Implementation**
Note: This figure visualizes the estimated degree of racially polarized voting over time in Oakland and San Francisco. The estimates from non-RCV elections are represented in gray, whereas those from RCV elections are shown in purple. The dots are weighted estimates and the triangles are unweighted estimates, where the weight is based on the proportion of voters in each cluster (in RCV elections) and voters supporting for each candidate (in non-RCV elections). The estimates are based on 50000 draws from the MCMC where the first 10000 draws are discarded as burn-in.

Next, we perform a descriptive analysis on the association between the RCV implementation and the degree of racially polarized voting. Specifically, we regress the posterior mean estimate of racially polarized voting on the dummy variable for RCV elections, dummy variable for San Francisco, and a discrete variable for year. Table 7 reports our results for both weighted and unweighted estimates of racial polarization. It shows that RCV elections are negatively associated with the level of racially polarized voting with the significance level of 0.01. While we cannot draw any causal conclusion from the above regression analyses, it at least demonstrates that there exists variation that we can further explore in more rigid causal analyses.

7 Concluding Remarks

In this article, we have attempted to provide the most direct evidence for the moderation hypothesis in the ranked-choice voting literature. While our data collection is so far limited to only two cities of Oakland and

	Ordinary Least Squares	
	Weighted	Unweighted
RCV	-0.27**	-0.14**
	(0.07)	(0.04)
San Francisco	0.03	-0.002
	(0.04)	(0.021)
Year	0.02*	0.01*
	(0.01)	(0.00)
N	12	12

Table 7: **Regression Analysis of Estimated Racially Polarized Voting**
Note: This table the results for regression analysis (OLS) where the posterior mean of racially polarized voting is regressed on the dummy variable for ranked-choice voting, dummy variable for San Francisco, and a discrete variable for year. * $p < 0.05$; ** $p < 0.01$.

San Francisco, we demonstrated that RCV elections are negatively associated with the degree of racially polarized voting. In our future research, we will perform a causal analysis based on the TSCS data by collecting more treated and control contests in the Bay Area.

Specifically, we expand the scope of the analysis by collecting precinct returns and cast vote records from Berkeley and San Leandro mayoral elections before and after the RCV implementation and other jurisdictions from the Bay Area in California which have never implemented RCV elections. Specifically, we will include mayoral elections in the cities of San Jose, Fremont, Santa Rosa, Hayward, Concord, Santa Clara, Vallejo, Fairfield, Richmond, Antoich, Daly City, San Mateo, Vacaville, and Livermore. We choose these Bay Area cities because they have similar population sizes as the cities that have implemented RCV elections.

Along with expanding the range of contests to analyze, we will delve into the causal mechanism of the RCV implementation and account for possible confounding factors. As discussed in Section 4, the implementation of RCV is strongly influenced by the level of minority descriptive representation and electoral environments in jurisdictions. To be able to compare similar jurisdictions with and without (the history of) RCV implementation, it is critical to fully understand where RCV could have been adopted and implemented. Despite the limited scope condition, we hope that this article shelds a new light on the relationship between RCV elections and racially polarized voting.

References

Alvarez, R Michael, Thad E Hall and Ines Levin. 2018. "Low-Information Voting: Evidence From Instant-Runoff Elections." *American Politics Research* 46(6):1012–1038.

Alvo, Mayer and Philip LH Yu. 2014. *Statistical methods for ranking data*. Springer.

Ansolabehere, Stephen, Nathaniel Persily and Charles Stewart III. 2009. "Race, religion, and vote choice in the 2008 election: implications for the future of the Voting Rights Act." *Harv. L. Rev.* 123:1385.

Ansolabehere, Stephen, Nathaniel Persily and Charles Stewart III. 2012. "Regional Differences in Racial Polarization in the 2012 Presidential Election: Implications for the Constitutionality of Section 5 of the Voting Rights Act." *Harv. L. Rev. F.* 126:205.

Atsusaka, Yuki. 2019. "A Framework for Analyzing Rank Data in Political Science." *Working Paper* .

Brace, Kimball, Lisa Handley, Richard G. Niemi and Harold W. Stanley. 1995. "Minority Turnout and the Creation of Majority-Minority Districts." 23(2):190–203.

Bratton, Michael and Mwangi S Kimenyi. 2008. "Voting in Kenya: putting ethnicity in perspective." *Journal of Eastern African Studies* 2(2):272–289.

Bratton, Michael, Ravi Bhavnani and Tse-Hsin Chen. 2012. "Voting intentions in Africa: ethnic, economic or partisan?" *Commonwealth & Comparative Politics* 50(1):27–52.

Brooks, Stephen P and Andrew Gelman. 1998. "General methods for monitoring convergence of iterative simulations." *Journal of computational and graphical statistics* 7(4):434–455.

Brown, Graham K. 2005. "Playing the (non) ethnic card: The electoral system and ethnic voting patterns in Malaysia." *Ethnopolitics* 4(4):429–445.

Bullock III, Charles S and Bruce A Campbell. 1984. "Racist or racial voting in the 1981 Atlanta municipal elections." *Urban Affairs Quarterly* 20(2):149–164.

Burnett, Craig M and Vladimir Kogan. 2015. "Ballot (and voter)exhaustion under Instant Runoff Voting: An examination of four ranked-choice elections." *Electoral Studies* 37:41–49.

Cain, Bruce E and Emily R Zhang. 2016. "Blurred Lines: Conjoined Polarization and Voting Rights." *Ohio St. LJ* 77:867.

Cain, Bruce E., Karin Mac Donald and Michael McDonald. 2005. From Equality to Fairness:: The Path of Political Reform since Baker v. Carr. In *Party Lines*, ed. Bruce E. Cain and Thomas E. Mann. Competition, Partisanship, and Congressional Redistricting Brookings Institution Press pp. 6–30.

Coakley, John. 2009. "The political consequences of the electoral system in Northern Ireland." *Irish political studies* 24(3):253–284.

Coakley, John and Jon Fraenkel. 2009. "Do preference transfers assist moderates in deeply divided societies? Evidence from Northern Ireland and Fiji.".

Coakley, John and Jon Fraenkel. 2017. "The Ethnic Implications of Preferential Voting." *Government and Opposition* 52(4):671–697.

Collet, Christian. 2005. "Bloc voting, polarization, and the panethnic hypothesis: The case of Little Saigon." *The Journal of Politics* 67(3):907–933.

Diaconis, Persi. 1988. "Group representations in probability and statistics." *Lecture notes-monograph series* 11:i–192.

Donovan, Todd, Caroline Tolbert and Kellen Gracey. 2016. "Campaign civility under preferential and plurality voting." *Electoral Studies* 42:157–163.

Elmendorf, Christopher S, Kevin M Quinn and Marisa A Abrajano. 2016. "Racially Polarized Voting." *U. Chi. L. Rev.* 83:587.

Engstrom, Richard L. 2015. "The Elephant in the Room: NAMUDNO, Shelby County, and Racially Polarized Voting." *Transatlantica. Revue détudes américaines. American Studies Journal* (1).

Fligner, Michael A and Joseph S Verducci. 1993. *Probability models and statistical analyses for ranking data*. Vol. 80 Springer.

Fraenkel, Jon. 2015. "Equality of Rights for Every Civilised Man South of the Zambezi: Electoral Engineering in Southern Rhodesia, 1957–65." *Journal of Southern African Studies* 41(6):1167–1180.

Fraenkel, Jon and Bernard Grofman. 2004. "A neo-Downsian model of the alternative vote as a mechanism for mitigating ethnic conflict in plural societies." *Public Choice* 121(3-4):487–506.

Fraenkel, Jon and Bernard Grofman. 2006*a*. "Does the alternative vote foster moderation in ethnically divided societies? The case of Fiji." *Comparative Political Studies* 39(5):623–651.

Fraenkel, Jon and Bernard Grofman. 2006*b*. "The failure of the alternative vote as a tool for ethnic moderation in Fiji: A rejoinder to Horowitz." *Comparative political studies* 39(5):663–666.

Fraenkel, Jon and Bernard Grofman. 2007. "The merits of Neo-Downsian modeling of the alternative vote: A reply to Horowitz." *Public choice* 133(1-2):1–11.

Frey, William H. 2018. *Diversity explosion: How new racial demographics are remaking America*. Brookings Institution Press.

Gerber, Elisabeth R and Jeffrey B Lewis. 2004. "Beyond the median: Voter preferences, district heterogeneity, and political representation." *Journal of Political Economy* 112(6):1364–1383.

Gormley, Isobel Claire and Thomas Brendan Murphy. 2008. "Exploring voting blocs within the Irish electorate: A mixture modeling approach." *Journal of the American Statistical Association* 103(483):1014–1027.

Gormley, Isobel Claire, Thomas Brendan Murphy et al. 2008. "A mixture of experts model for rank data with applications in election studies." *The Annals of Applied Statistics* 2(4):1452–1477.

Greiner, D James. 2011. "Re-Solidifying Racial Bloc Voting: Empirics and Legal Doctrine in the Melting Pot." *Ind. LJ* 86:447.

Greiner, D James and Kevin M Quinn. 2009. "R$\times$ C ecological inference: bounds, correlations, flexibility and transparency of assumptions." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 172(1):67–81.

Greiner, D James and Kevin M Quinn. 2010. "Exit polling and racial bloc voting: Combining individual-level and R x C ecological data." *The Annals of Applied Statistics* pp. 1774–1796.

Grofman, Bernard. 1991. "Multivariate methods and the analysis of racially polarized voting: pitfalls in the use of social science by the courts." *Social Science Quarterly* 72(4):826–833.

Grofman, Bernard and Michael Migalski. 1988. "Estimating the extent of racially polarized voting in multicandidate contests." *Sociological Methods & Research* 16(4):427–453.

Grofman, Bernard and Scott L Feld. 2004. "If you like the alternative vote (aka the instant runoff), then you ought to know about the Coombs rule." *Electoral studies* 23(4):641–659.

Hajnal, Zoltan and Jessica Trounstine. 2014. "What underlies urban politics? Race, class, ideology, partisanship, and the urban vote." *Urban Affairs Review* 50(1):63–99.

Hajnal, Zoltan L. 2009. "Who loses in American democracy? A count of votes demonstrates the limited representation of African Americans." *American Political Science Review* 103(1):37–57.

Hasen, Richard L. 2018. "Race or Party, Race as Party, or Party All the Time: Three Uneasy Approaches to Conjoined Polarization in Redistricting and Voting Cases." *Wm. & Mary L. Rev.* 59:1837.

Herron, Michael C and Jeffrey B Lewis. 2007. "Did Ralph Nader spoil Al Gore's presidential bid? A ballot-level study of green and reform party voters in the 2000 presidential election." *Quarterly Journal of Political Science* 2(3):205–227.

Horowitz, Donald L. 1991*a*. *A democratic South Africa? Constitutional engineering in a divided society*. Berkeley: University of California Press.

Horowitz, Donald L. 1991*b*. Ethnic conflict management for policymakers. In *Conflict and peacemaking in multiethnic societies*, ed. Joseph V Montville. Lexington, MA: Lexington Books pp. 115–130.

Horowitz, Donald L. 1991*c*. Making moderation pay: The comparative politics of ethnic conflict management. In *Conflict and peacemaking in multiethnic societies*, ed. Joseph V Montville. Lexington, MA: Lexington Books pp. 451–476.

Horowitz, Donald L. 2000. *Ethnic groups in conflict, updated edition with a new preface*. Univ of California Press.

Horowitz, Donald L. 2004. "The alternative vote and interethnic moderation: A reply to Fraenkel and Grofman." *Public Choice* 121(3-4):507–517.

Horowitz, Donald L. 2006. "Strategy takes a holiday: Fraenkel and Grofman on the alternative vote." *Comparative Political Studies* 39(5):652–662.

Horowitz, Donald L. 2007. "Where have all the parties gone? Fraenkel and Grofman on the alternative vote–yet again." *Public Choice* 133(1-2):13–23.

Huber, John D. 2012. "Measuring ethnic voting: Do proportional electoral laws politicize ethnicity?" *American Journal of Political Science* 56(4):986–1001.

Imai, Kosuke, In Song Kim and Erik Wang. 2018. "Matching methods for causal inference with time-series cross-section data." *Princeton University* 1.

Issacharoff, Samuel. 1992. "Polarized voting and the political process: The transformation of voting rights jurisprudence." *Michigan Law Review* 90(7):1833–1891.

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.

John, Sarah, Haley Smith 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.

Kimball, David C and Joseph Anthony. 2016. Voter participation with ranked choice voting in the United States. In *annual meeting of the American Political Science Association, Philadelphia*.

Kimball, David C and Joseph Anthony. 2018. Ranked Choice Voting: A Different Way of Casting and Counting Votes. In *Changing How America Votes*, ed. Todd Donovan. Lanham, MD: Rowman Littlefield Publishers chapter 8, pp. 100–112.

King, Gary. 1997. *A solution to the ecological inference problem: Reconstructing individual behavior from aggregate data*. Princeton University Press.

Kuriwaki, Shiro. 2019. "The Administration of Cast Vote Records in U.S. States." *Working Paper* .

Lau, Olivia, Ryan T Moore and Michael Kellermann. 2007. "eiPack: R× C Ecological Inference and Higher-Dimension Data Management." *New Functions for Multivariate Analysis* 7(1):43.

Liu, Qinghua, Marta Crispino, Ida Scheel, Valeria Vitelli and Arnoldo Frigessi. 2018. "Model-Based Learning from Preference Data." *Annual Review of Statistics and Its Application* (0).

Lupia, Arthur and Kenneth McCue. 1990. "Why the 1980s Measures of Racially Polarized Voting Are Inadequate for the 1990s." *Law & Policy* 12(4):353–387.

Mallows, Colin L. 1957. "Non-null ranking models. I." *Biometrika* 44(1/2):114–130.

Maloy, JS. 2019. *Smarter Ballots: Electoral Realism and Reform*. Springer.

Marden, John I. 1995. *Analyzing and modeling rank data*. Chapman and Hall/CRC.

McDaniel, Jason. 2018. "Does More Choice Lead to Reduced Racially Polarized Voting? Assessing the Impact of Ranked-Choice Voting in Mayoral Elections." *California Journal of Politics and Policy* 10(2).

McDaniel, Jason A. 2016. "Writing the rules to rank the candidates: Examining the impact of instant-runoff voting on racial group turnout in San Francisco mayoral elections." *Journal of Urban Affairs* 38(3):387–408.

Mitchell, Paul. 2014. "The single transferable vote and ethnic conflict: the evidence from Northern Ireland." *Electoral Studies* 33:246–257.

Park, Won-ho, Michael J Hanmer and Daniel R Biggers. 2014. "Ecological inference under unfavorable conditions: straight and split-ticket voting in diverse settings and small samples." *Electoral Studies* 36:192–203.

Plescia, Carolina and Lorenzo De Sio. 2018. "An evaluation of the performance and suitability of R× C methods for ecological inference with known true values." *Quality & quantity* 52(2):669–683.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
URL: *http://www.R-project.org/*

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

Reilly, Benjamin. 2018. "Centripetalism and electoral moderation in established democracies." *Nationalism and Ethnic Politics* 24(2):201–221.

Robb, Denise. 2011. *The effect of instant runoff voting on democracy*. University of California, Irvine.

Rosen, Ori, Wenxin Jiang, Gary King and Martin A Tanner. 2001. "Bayesian and frequentist inference for ecological inference: The R$\times$ C case." *Statistica Neerlandica* 55(2):134–156.

Schultz, David and Kristi Rendahl. 2010. "Evaluating ranked choice voting in the 2009 Minneapolis elections: A report for the Minneapolis elections department.".

Shamir, Michal and Asher Arian. 1982. "The ethnic vote in Israel's 1981 elections." *Electoral Studies* 1(3):315–331.

Sisk, Timothy D. 1996. *Power sharing and international mediation in ethnic conflicts*. US Institute of Peace Press.

Taebel, Delbert. 1978. "Minority Representation on City Councils: The Impact of Structure on Blacks and Hispanics." 59(1):142–152.
URL: *https://www.jstor.org/stable/42859874*

Tam, Wendy K. 1995. "Asians – A monolithic voting bloc?" *Political Behavior* 17(2):223–249.

Vedlitz, Arnold and Charles A. Johnson. 1982. "Community Racial Segregation, Electoral Structure, and Minority Representation." 63(4):729–736.
URL: *https://www.jstor.org/stable/42861463*

Vitelli, Valeria, Øystein Sørensen, Marta Crispino, Arnoldo Frigessi and Elja Arjas. 2018. "Probabilistic preference learning with the Mallows rank model." *Journal of Machine Learning Research* 18(158):1–49.

Appendices

A Details of San Francisco and Oakland Mayoral Elections

In this section, we report details of San Francisco and Oakland mayoral elections we analyze in this article. The next table shows the number of rounds, number of candidates, and information about incumbents in all contests analyzed in this article.

City	Year	RCV	#Rounds	#Candidates	Incumbent
Oakland	2018	✓	1	10+1	reelected
Oakland	2014	✓	16	16+1	defeated
Oakland	2010	✓	10	10+1	open seat
Oakland	2006		1	6+1	open seat
Oakland	2002		1	2	reelected
Oakland	1998		1	11	open seat
San Francisco	2019	✓	1	6+1	reelected
San Francisco	2018	✓	9	8+1	reelected*
San Francisco	2015	✓	1	6+1	reelected
San Francisco	2011	✓	12	16+1	reelected*
San Francisco	2007	✓	1	12+6	reelected
San Francisco	2003		2	9+1	open seat
San Francisco	1999		2	14+4	reelected
San Francisco	1995		2	8	defeated

Table 8: **Details in Mayoral Elections in San Francisco and Oakland**
Note: This table lists the number of rounds, number of candidates, and incumbency status in all elections analyzed here. When there is any write-in candidate, she or he is denoted by "+" in Column (5).

B Descriptive Statistics

In this section, we explore descriptive statistics in our data.

City	Year	White	Asian	Hispanic	Black
Oakland	2018	0.32	0.12	0.07	0.24
Oakland	2014	0.34	0.12	0.07	0.23
Oakland	2010	0.40	0.13	0.09	0.29
Oakland	2006	0.43	0.14	0.09	0.30
Oakland	2002	0.41	0.11	0.08	0.36
Oakland	1998	0.41	0.11	0.08	0.36
San Francisco	2019	0.49	0.30	0.13	0.05
San Francisco	2018	0.49	0.30	0.13	0.05
San Francisco	2015	0.45	0.28	0.17	0.07
San Francisco	2011	0.44	0.31	0.15	0.07
San Francisco	2007	0.44	0.31	0.15	0.07
San Francisco	2003	0.53	0.27	0.11	0.06
San Francisco	1999	0.51	0.28	0.11	0.06
San Francisco	1995	0.54	0.23	0.12	0.10

Table 9: **Descriptive Statistics in San Francisco and Oakland**
Note: This table lists the proportion of white, black, asian, and hispanic voters in analyzed contests.

C Notes on Ecological Inference Estimates

This section describes technical details and robustness of our ecological inference estimates. There are two major classes of ecological inference (EI) models currently available for multiple racial groups and candidates (i.e., R $\times$ C models): the Multinomial-Dirichlet model (Rosen et al., 2001) and the GQ model (Greiner and Quinn, 2009, 2010).[12] In this article, we employed the former model due to our limited computation power. However, future extensions must replicate our results based on the second model.

Moreover, due to computational limit, we had to limit the length of the Markov chains that explore the space of the posterior distribution up to 50000 and drop 10000 first draws as burn-in. As diagnostic statistics show (Brooks and Gelman, 1998), the convergence of the chains seems to be insufficient to draw more decisive conclusions about our estimates. All of these limitations mean that we must be cautious about reading our results.

[12] For a validation study for both models, see Plescia and De Sio (2018).