UCL Centre on US Politics
Working Paper Series in American Politics

Working Paper No. 2022-1

Electoral Institutions and Substantive Representation in Local Politics: The Effects of Ranked Choice Voting

Arjun Vishwanath
(Harvard University)

avishwanath1@g.harvard.edu

Suggested Citation: Vishwanath, Arjun. 2022. "Electoral Institutions and Substantive Representation in Local Politics: The Effects of Ranked Choice Voting." CUSP Working Paper No. 2022-1. London: University College London

Electoral Institutions and Substantive Representation in Local Politics: The Effects of Ranked Choice Voting*

Arjun Vishwanath†
Harvard University

July 14, 2022

Abstract

While studies of substantive representation in local politics have advanced considerably in the past decade, we know less about the effects of institutions on local representation. I explore the effect of an increasingly popular electoral institution—ranked choice voting (RCV)—on substantive representation. Examining a host of fiscal variables and the ideological composition of city councils, I find that the municipalities adopting RCV saw no changes in outcomes or representation. I also use roll call data to show that the behavior of legislators who served both before and after the use of RCV did not change and that intra-city responsiveness remained constant after implementation. Lastly, I find that RCV did not even induce ideological changes in the set of candidates who decided to run. I conclude that RCV does not produce the significant effects on representation for which its proponents hope.

Word count: 9,670

*Acknowledgments: I am grateful to the New America Foundation and Arnold Ventures for funding the research. Mirnes Kukic, Shelly Tsirulik, and Sophia Lee provided excellent research assistance. Chris Warshaw and Peter Bucchianeri graciously shared data for the analyses. Jim Snyder, Noah Daponte-Smith, Tyler Simko, and participants in the Electoral Reform Research Group gave helpful feedback. I am especially indebted to Jon Rogowski for detailed comments and advice.

†Ph.D. candidate, Department of Government; avishwanath1@g.harvard.edu.

Recently, political contestation has focused on the institutional structure of politics itself. Politicians and activists recognize that these rules—in particular, electoral rules—influence the composition of the government and, as a result, policy outcomes. As such, these institutions play a central role in shaping political representation. One electoral rule proposal that has gained traction among reformers in the past decade is ranked choice voting (RCV). Approximately two dozen states and cities have implemented it, including Maine, Alaska, and New York City in recent years. Proponents suggest that RCV enables voters to cast a sincere ballot rather than a strategic one. They also argue that it forces candidates to avoid negative campaigning and moderate their platforms as a result of the need to acquire second and third place votes (Richie 2004). These reformers conclude that RCV can resolve the frustrations of the two party system and empower the voices of all voters. As the Ranked Choice Voting Resource Center argues on its website in a section entitled 'Promoting Fair Representation', "legislatures elected by winner-take-all may lead to distortions in partisan representation, the entrenchment of incumbents in safe seats, regional polarization, and low representation of women and racial and ethnic minorities."[1] Mainstream voices echo reformers' arguments, as the *New York Times* Editorial Board did in its endorsement of the proposal for presidential primaries and New York City elections.[2] A *Washington Post* opinion piece even argues that RCV, among all electoral reforms, is "most likely to save democracy."[3]

However, there is scant empirical evidence to support these ambitious claims, and RCV's representational consequences in particular remain unexplored in political science. Literature on the effects of other institutions on representation—especially in local politics, where RCV

[1] https://www.rcvresources.org/why-adopt-rcv. Accessed June 28, 2021.

[2] The Editorial Board. "The Primaries Are Just Dumb." The *New York Times*, 26 February, 2020.
The Editorial Board. "New Yorkers Have a Chance to Remake How They Vote." The *New York Times*, 15 October, 2020.

[3] Edward B. Foley. "Can Alaska save democracy?" The *Washington Post*, 10 February, 2022.

has largely been adopted to date—provides little guidance to validate proponents' claims. These studies suggest that institutions, such as term limits or direct democratic institutions, have few mediating effects on policy outcomes (De Benedictis-Kessner and Warshaw 2016; Tausanovitch and Warshaw 2014). While the study of local representation has advanced considerably in the past decade, we still know little about which mechanisms affect this representation.

In this paper, I study the effect of RCV on substantive policy representation in local politics.[4] The main claim I explore is whether RCV induces changes in fiscal outcomes and the ideological composition of city councils. I find no effect of RCV on policy outcomes or city council ideology on their own terms or relative to voter preferences in the adopting cities. I also use roll call data from local governments in San Francisco and Oakland to determine whether there are within-legislator shifts in voting behavior following RCV adoption and whether within-city ideological responsiveness across city districts is affected; again, I find no effect from the reform. Finally, I address the more modest claim of advocates that RCV induces changes in the candidate pool (regardless of who wins). Here, I look at shifts in the distribution of candidates for city council before and after adoption and again find minimal evidence of any shifts.

My findings are largely consistent with existing literature that has found minimal effects of institutions on local politics. Furthermore, my multi-method approach enables me to address different theories of local politics. For example, the null findings on fiscal outcomes could be explained by the legal and political constraints that municipalities face (Peterson 1981), but I also find null effects in the ideological composition of city councils. In sum, it appears that RCV, like many other electoral reforms to gain in popularity in recent decades, does not meaningfully affect American politics.

[4] I do not consider single-transferable vote (STV) systems like the one in Cambridge, MA.

1 Institutions and Representation in Local Elections

Recent studies of representation in local politics have found a correspondence between public opinion and policy. In the most comprehensive treatment of the question, Tausanovitch and Warshaw (2014) construct estimates of municipal public opinion and compare these to cities' policy outcomes. They find that more conservative cities have lower taxes and expenditures, suggesting some degree of responsiveness. Einstein and Kogan (2016) also find a correlation between more Democratic cities and higher levels of spending, taxation, and intergovernmental transfers.

However, there has been little work to date on how institutions *influence* the caliber of local representation. Extant literature on local representation has focused on three relationships: the effect of mayoral partisanship on fiscal outcomes, the effect of economic conditions on local voting behavior, and the effect of voters' policy preferences on the ideology of local government and policy outcomes. In each, the evidence as to whether institutions play a mediating role is mixed and inconclusive at best.

The most studied institutional relationship is that of officeholder partisanship on policy outcomes. Ferreira and Gyourko (2009) find that partisan control of the mayoralty does not affect fiscal outcomes. They suggest that the relationship is mediated by the homogeneity of city populations and "Tiebout competition" between cities in the same metropolitan area, but access to local news does not impact the relationship. Gerber and Hopkins (2011) instead find that partisan control does matter: Republican mayors allocate a larger proportion of the local budget to police and fire spending than Democrats do. These effects are produced by partisan elections; in cities without partisan elections, they find no difference in spending allocations. More recently, De Benedictis-Kessner and Warshaw (2016) find that Democratic mayors spend more per capita, enact higher taxes per capita, and issue larger amount of debt than their Republican counterparts do. Contra Gerber and Hopkins, though, they find *no*

difference between partisan and nonpartisan elections. They also find no differences in representation between cities with a council-manager form of government and those with strong mayor forms. Approaching the question from a different angle, Einstein and Glick (2018) survey mayors and find that Democrats are more likely to value redistribution and issue statements about it. They too find no effect of a strong mayor system on this relationship.

Turning to the relationship between economic outcomes and local voting, Hopkins and Pettingill (2018) find that mayors receive an electoral boost when their cities' unemployment rate is lower than the national rate. However, the degree of retrospection is unaffected by institutions such as partisan nominations and balloting, electoral timing, the form of government, and term limits (although it is affected by media access). By contrast, Berry and Howell (2007) only find responsiveness in school board elections to student performance under highly contingent circumstances; in two of the three years they study, they find no responsiveness. Furthermore, this relationship is unaffected by the presence of partisan elections or levels of legislative compensation.[5]

Finally, there are mixed findings on the factors that influence local policy responsiveness to public opinion. Tausanovitch and Warshaw (2014) find that policy responsiveness to public opinion is unaffected by the form of government, the existence of direct democracy, partisan elections, term limits, and at-large elections. That said, Dynes, Hartney, and Hayes (2021) extend Tausanovitch and Warshaw's analysis to look at the timing of local elections and find that off-cycle elections reduce responsiveness to public opinion.

Across all three relationships, the evidence that electoral institutions systematically affect representation is mixed at best. While this does not necessarily mean RCV cannot affect local politics, there is little precedent for the type of institutionally-driven systemic change

[5]De Benedictis-Kessner and Warshaw (2020a) find that local voting is influenced by presidential partisanship, although they do not test whether it is influenced by the partisanship of the local officeholder. That said, unlike Hopkins and Pettingill, they find that the economic voting relationship is not mediated by media coverage.

that advocates predict.

2 Candidate and Voter Behavior Under RCV

In hopes of reconfiguring the existing American political system, electoral reformers have sought to alter the way Americans vote using a method called ranked choice voting. RCV enables voters to cast a "preference" ballot rather than indicating a single choice under the first-past-the-post (FPTP) method. The voter ranks one candidate as her first choice, another as her second, and so on. The first place choices are aggregated, and if the leader among first place ballots does not receive a majority, the candidate with the fewest first-place votes is eliminated, and the ballots for that candidate are reallocated to their next highest ranked choices. If a voter does not a complete a ballot or all of her choices are eliminated, the ballot is exhausted and removed from the denominator. This process is repeated until one candidate receives a majority.

RCV was first adopted in San Francisco in 2004, and it has since spread to a number of cities, including New York City starting in 2021. While RCV has largely been implemented by cities to date, a growing reform movement has attempted to install it at the state level as well. Maine implemented it in its 2018 federal and state elections, several states used the method in the 2020 Democratic presidential primaries, and Alaska began using a variant of it in 2022. Its rise can be attributed to 'good government' reformers who argue that RCV eliminates two barriers to effective representation stemming from FPTP.

First, they argue that FPTP does not reflect the sincere preferences of voters because FPTP enables candidates to win without majority support and requires voters to cast ballots strategically in making this plurality choice (Richie 2004). By comparison, under RCV, voters cast a preference ballot as opposed to selecting a single candidate. Thus, in a race between a Democrat, Republican, and a Libertarian candidate, a libertarian-leaning con-

servative voter may rank the Libertarian candidate first, the Republican second, and the Democrat third. Her vote will not be wasted if the Libertarian is eliminated, and so her sincere preference rankings are considered in the balloting.[6]

While this rationale appears attractive, the empirical evidence for it is ambiguous. On the one hand, voters *do* express clearly identifiable preferences in ranked choice elections. Alvarez, Hall, and Levin (2018) examine the latent factors underlying the preference rankings of ballots and find that, especially in partisan races, a few dimensions explain much of the variation in voter preferences.[7] However, survey research indicates that RCV rarely changes the election winner, and it does not improve confidence in elections or democracy; in fact, strong majorities prefer plurality voting to RCV even after attempting to use RCV (Nielson 2017). Additionally, most of the empirical evidence evaluating changes in campaign strategy and voter behavior relies on qualitative assessments (Richie 2004; Robb 2011), survey results of voters' attitudes towards the system (Donovan, Tolbert, and Gracey 2016; Nielson 2017), or one-off case studies (Alvarez, Hall, and Levin 2018). As such, the evidence produced by these studies cannot fully address the claims made by proponents of RCV.

Second, RCV advocates argue that FPTP encourages candidate polarization and negative campaigning in order to appear as the "least bad" candidate. They assert that RCV reduces negative campaigning and leads to candidate moderation since candidates must garner second- and third-choice preferences from voters outside their core base (Richie 2004). Surveys indicate that voters in RCV systems express higher levels of satisfaction with the way candidates have conducted themselves and say that candidates are less likely criticize each other (Donovan, Tolbert, and Gracey 2016). Robb (2011) examines the content of mail-

[6]The first problem is partially mitigated in two-round runoffs where no candidate receives a majority. However, the second problem still persists for those elections. Additionally, turnout in the second round of runoffs is lower than the votes lost due to ballot exhaustion under RCV (Kimball and Anthony 2016).

[7]They find in nonpartisan races that candidate traits tend to prevail, though.

ers in San Francisco and finds that negative campaigning decreased following RCV adoption. However, no research to date has examined changes in candidate moderation following RCV adoption.

Of course, there are also reasons why RCV may not improve representation or may even worsen representation. The most notable is that RCV fails to produce "monotonicity" (Fishburn and Brams 1983). Put simply, this failure means that if the ballots of certain voters were altered to push the would-be winner higher (holding all other ballots constant), then a *different candidate* might actually win the election.[8] This phenomenon occurs because altering the rankings can result in the winner facing a different candidate in a later round to whom they lose. Ornstein and Norman (2014) show that monotonicity fails between 15% and 50% of the time, depending on the distribution of the electorate and candidates. In addition, Kilgour, Grégoire, and Foley (2020) show that different degrees of ballot truncation (i.e., ranking fewer candidates than the total number running) can produce different winners even with the same latent preferences. This is especially true when ballot truncation is forced (e.g., San Francisco voters can rank at most three candidates) but can happen even when voters voluntarily submit incomplete ballots.

More broadly, RCV fails to ensure that a "Condorcet winner" will win the election (i.e., the candidate who beats every other candidate in a pairwise matchup). Though every Bay Area election held under RCV has produced a Condorcet winner (whereas FPTP would have produced some non-Condorcet winners), the 2009 Burlington, VT mayoral race did not. There, the Democrat was the Condorcet winner but was eliminated with three candidates left, and the Progressive candidate prevailed over the Republican in the final round.[9] This failure to select a Condorcet winner was a factor in Burlington's decision to repeal RCV

[8]The reverse is also possible—a loser could win if they were ranked lower by some voters.

[9]https://www.fairvote.org/every_rcv_election_in_the_bay_area_so_far_has_produced_condorcet_winners

(Ornstein and Norman 2014).

Leaving aside theoretical problems, the complexity of RCV can produce difficulties for voters. Burnett and Kogan (2015) find high levels of ballot exhaustion. This can take place due to forced truncation or due to improper or incomplete ballots—for example, a ballot may rank the same candidate three times or may leave ranking spots blank. Thus, the victor could be chosen with less than a majority of total ballots cast in the race, undermining a key argument of RCV advocates. Burnett and Kogan also note that one of the supposed upsides of RCV—that it allows voters to avoid the need for strategic voting—is not actually valid. If a voter casts a sincere ranking of a few candidates based on this understanding, her ballot may be exhausted early. Worse yet, the assumption that RCV permits sincere voting is also inaccurate—there are numerous scenarios under which voters ought to vote strategically even under RCV, and the strategic considerations are far more complex than those in FPTP.[10]

3 How RCV Might Affect Representation

While much of the RCV literature has focused on its campaign effects, there is far less work on its legislative and representational effects, and what little work has been done suggests mixed effects. On the descriptive side, John, Smith, and Zack (2018) find that the proportion of minority candidates and number of female winners increased in the four Bay Area cities employing RCV. However, McDaniel (2016) finds that RCV reduced black

[10]Consider an election where a voter has the preferences $A \succ B \succ C$. There are two voters with her same preferences, three voters with $C \succ A \succ B$, and two more voters with $B \succ C \succ A$. If our voter votes sincerely, A and C will advance to the next round, and C will garner 5 votes to A's 3. If, however, our voter casts an insincere ballot in which she ranks B ahead of A, then B and C will advance to the next round, and B will garner 5 votes to C's 3—thus yielding a better outcome for our voter. However, such calculations are nearly impossible for a well-informed voter, let alone the typical one.

turnout in San Francisco, and Santucci and Scott (2021) find little evidence that minorities and women are encouraged to run after learning about RCV. On the substantive side, almost no work exists, although there is some evidence that it alters legislative coalitions (Santucci 2018). Furthermore, no research to date has examined the question of how RCV itself affects policy outcomes or substantive representation, which is the question I aim to answer.

There are a few ways in which RCV could enhance representation. First, it can put more representative candidates into office holding constant voters' and candidates' ideological positions. This mechanism corresponds to the argument made by RCV proponents that FPTP inefficiently accounts for voter preferences. There is ample evidence that voters employ strategic considerations (Hall and Snyder 2015), and so the switch to RCV may induce an increase in sincere voting. Voters may rank third party candidates or other primary contenders higher whom they previously judged to be non-viable. In doing so, those candidates may garner enough votes to win. Alternatively, RCV could reallocate the votes of current third-party voters in a way that swings an election from the Republican to the Democrat or vice-versa. For example, in the Maine gubernatorial elections in 2010 and 2014, a left-leaning independent took enough votes away from the Democrat to enable conservative Republican Paul LePage. On the other side, the Libertarian candidate took enough votes away from Sen. David Perdue in the 2020 Georgia Senate race to prevent him from winning outright, and Perdue proceeded to lose in a lower-turnout runoff. That said, Abramson et al. (2010) argue that strategic voting remains at comparable levels in a cross-national study of FPTP versus PR systems. In Appendix Section A, I show that RCV need not improve representation (measured in terms of the likelihood of selecting the median-preferred voter) compared to FPTP. Instead, the likelihood of identifying the optimal candidate varies based on the distribution of voters and number of candidates running.

Second, RCV could push legislators' positions towards the median voter. This could take place due to the reward for cooperative campaign behavior, or it could simply be a

result of the requirement that a candidate be supported by a majority of voters under RCV. Insofar as legislators move to the median by proposing and adopting policies that the median favors, we should see policy outcomes such as taxation and spending levels shift towards the median's preference. Some formal models suggest that, under certain conditions, RCV can play such a moderating influence on legislator behavior (Buisseret and Prato 2022; Dellis, Gauthier-Belzile, and Oak 2017).

A final, more modest possibility is that RCV could encourage more representative *candidates* to run who otherwise would not have run under FPTP elections. If this were the case, it would align with other work finding that most of the polarization in Congress can be attributed to lowered incentives for moderates to run (Hall 2019). Reformers have sought to rectify this by changing election methods with mixed results (e.g., Kousser, Phillips, and Shor 2018). Of course, a more moderate candidate pool that does not alter the winning candidate set may not seem to improve representation. However, in giving voice to different perspectives, citizens may feel that a wider range of views are represented in electoral debates.

4 Measuring Changes in Municipal Outcomes

I evaluate the impact of RCV on policy and ideological outcomes using a number of approaches. In keeping with existing literature estimating the effect of institutions on municipal spending (De Benedictis-Kessner and Warshaw 2016; De Benedictis-Kessner and Warshaw 2020b; Gerber and Hopkins 2011), I examine changes across sixteen municipal policy outcomes from 1996 to 2019 from US Census Bureau data (Pierson, Hand, and Thompson 2015): total revenues, general revenues, total taxes, property taxes, general sales tax, direct expenditures, charges and miscellaneous revenue, outstanding debt, parks and recreation spending, highway spending, sewerage spending, fire protection spending, police protection

spending, public welfare spending, library spending, and total utilities spending.[11] Where possible, I employ measures of direct spending that exclude intergovernmental transfers in order to assess the direct impact of the policy change on spending outcomes.

Measuring the ideological composition of city councils is slightly trickier. While there are now commonly accepted metrics of ideology for state legislators (Shor and McCarty 2011; Bonica 2014), there is no equivalent for local officials. In this article, I use campaign contribution-driven estimates of ideology, or CFscores (Bonica 2014). Unlike many studies that employ CFscores in federal and state politics, Bonica's database does not contain candidate scores for most local politicians. As a result, I use the candidates' *contributor scores* to estimate their ideology (that is, their ideal points as estimated from the other candidates to whom they have contributed) from the dataset of local election returns collected by Warshaw et al. (2022) from 1996 to 2021. I have supplemented this database by manually searching for all elected officials in RCV-adopting jurisdictions since 1990 and matching them to Bonica's database. I then aggregate the data at the city council-year level based on the timing of the electoral terms in each city to construct an estimate of the council ideology in a given year.[12]

As a measure of internal validity, Bonica (2014, 373) notes that candidate and contributor CFscores of those who run are highly correlated, suggesting these contributor scores are a reasonable proxy. While contributor scores have been used in this way before (De Benedictis-Kessner and Warshaw 2020b), they could be biased if a candidate strategically contributes to politicians for her own political advancement. In that case, contributions would not represent a sincere expression of political belief or a consumptive act. However, the fact that the vast majority of local politicians never advance past the municipal level suggests

[11] I exclude a seventeenth variable, education spending, in the main models since there is data for only one treated city (Portland), but it is included in the appendix analyses.

[12] I remove observations where there is only a CFscore for one member in the year or there are CFscores in that year for fewer than 25% of the maximum members who ever served on the council.

that strategic considerations are unlikely to skew their contribution behavior. Although they may contribute to political allies, it is more likely that they are doing so to signal ideological alignment in a manner akin to signalling one's ideology through roll call votes.[13]

These metrics enable us to explore changes in fiscal policy and the ideological composition of city governments. However, they are insufficient for answering questions about representation without an understanding of the demand for those changes. An increase in spending could be bad for representation if the public wants less spending or good for representation if the public wants more of it. Furthermore, null effects may mask changes in representation if, for example, the RCV-adopting municipalities that want more spending increase spending while those that want less spending decrease it.

To measure representation, I adapt a method from Rogers (2017) and Kilborn and Vishwanath (2022) to estimate representation. Those articles estimate state legislative representation by scaling legislative ideal points (from both roll-call- and contribution-based estimates) to state district-level ideology. In this case, I compare fiscal or ideological measures to each city's municipal ideology (Tausanovitch and Warshaw 2014). To do so, I fit a multivariate linear regression of the dependent variable in question (e.g., taxes per capita) on the city's mass liberalism in public opinion with year fixed effects for *all* cities. The fitted values from this regression provide the expected value of the dependent variable of a city with a certain ideology in a certain year. As such, these values can be thought of as "public demand" for the service. For example, the model suggests that a city as liberal as Oakland in 2002 would collect $1,141 in taxes per capita. Oakland actually collected $928 per capita, so this metric suggests that its tax revenue was lower than the level set by cities of Oakland's ideological bent. On the ideological side, the mean CFscore on the Oakland city council in

[13]Candidate CFscores have been critiqued for exhibiting low intraparty correlations with DW-NOMINATE scores (Barber 2022; Tausanovitch and Warshaw 2017). That said, Barber notes that this is not the case for state legislators, and this paper analyzes a different group altogether (local candidates using contributor scores rather than candidate scores).

2002 was -0.75. Cities whose electorates are of the same ideological bent as Oakland's had a mean CFscore of -1.25 for their city council, which suggests that Oakland was a bit more conservative than expected.[14] With this estimate of public demand, I calculate the representation metric as the absolute value of the difference between the two quantities ($213 for Oakland's tax per capita and 0.5 CFscore units for Oakland's city council in 2002).[15]

This method of estimating representation raises two potential concerns. The first is that the absolute difference may not be an unbiased estimate of the nature of representation. If cities systematically spend less than their citizens want, for example, this metric would fail to account for that. Since I am looking at changes in the metric after adopting RCV, however, I am concerned with *relative* differences in the metric over time. Furthermore, since the equation contains year fixed effects, it does not assume that this bias is constant over time. The second is that the single-dimensional measure of mass ideology may not correspond to the variation in preferences over each of the fiscal items I measure. For example, the measure may capture liberalism on tax policy but not on library spending. However, Tausanovitch and Warshaw (2014, 625-629) use a number of survey questions on both federal and local issues and find that opinion on local issues is one-dimensional and strongly correlated with views on federal issues. Furthermore, Figure 2 in Tausanovitch and Warshaw (2014, 612) shows that their ideology metric is closely correlated with taxation and spending levels in cities. The model here essentially re-estimates their regressions (adding year fixed effects) and tracks whether cities move closer or farther to nationwide levels of responsiveness.

One important factor in estimating RCV's effects is that the cities adopting RCV are highly dissimilar from those that do not. Table 1 shows the RCV-adopting cities that I

[14] This assumption is justified by other work shows that voting behavior in municipal governments is low-dimensional (Bucchianeri 2020).

[15] In mathematical terms, the representation metric is $D_{it} = |\hat{Y_{it}} - Y_{it}|$ where $\hat{Y_{it}} = \alpha_t + \beta C_i$. Y is the dependent variable for unit i in period t, C_i is the mass liberalism of the city, and α_t are year fixed effects.

Place	Year Adopted	Year Ended	Fiscal Analysis	City Council Analysis
San Francisco, CA	2004	Still Active	Yes	Yes
Burlington, VT	2006	2010	Yes	No
Aspen, CO	2007	2010	Yes	No
Cary, NC	2007	2011	Yes	Yes
Hendersonville, NC	2007	2011	No*	No
Takoma Park, MD	2007	Still Active	No*	No
Telluride, CO	2008	Still Active	No*	No
Minneapolis, MN	2009	Still Active	Yes	Yes
Berkeley, CA	2010	Still Active	Yes	Yes
Oakland, CA	2010	Still Active	Yes	Yes
San Leandro, CA	2010	Still Active	No*	Yes
Portland, ME	2011	Still Active	Yes	Yes
St. Paul, MN	2011	Still Active	Yes	Yes
Santa Fe, NM	2018	Still Active	No*	No
Las Cruces, NM	2019	Still Active	No	No*
St. Louis Park, MN	2019	Still Active	No	No*

Table 1: Table of cities analyzed. Starred cities are excluded from main models due to limited data in pre- and post-treatment periods but are included in appendix robustness checks.

analyze. To compare these cities with those that never adopted RCV, I collect demographic covariates from the 1990 and 2000 Censuses and supplement it with rolling 5 year ACS estimates for every year since 2009 (Manson et al. 2021).[16] I use the total population, percent male, percent that is white, black, Asian, and Hispanic, the percent with at least a college degree, the median income, the median home value, and the proportion of residents that are homeowners in the models. I also use two-party presidential vote in the most recent election at the county level. The RCV-adopting cities are more populous, more racially diverse, more educated, higher income, have higher home values, and are more likely to be renters relative to the cities that never adopted RCV. Perhaps most importantly, they are located in counties far more Democratic than the rest of the country (see Table B.1 for

[16]This data is collected at the census place level. I use the term "city" interchangeably with "census place" for the sake of interpretability. Since ACS data for 2021 is not yet available, I use the 2020 estimates for the 2021 city council models.

coefficients).

To produce a more appropriate comparison set, I weight the control units based on treatment history and these covariates. I construct the weights based on covariate balancing propensity scores in the 7 periods prior to treatment adoption.[17] I then estimate the average treatment effect on the treated units under a differences-in-differences framework using the `PanelMatch` package (Imai, Kim, and Wang 2021). Appendix Figures B.1 and B.2 show that the procedure substantially improves the covariate balance between the treatment and control groups. In Appendix Section D.1, I also show that my findings are robust to another possible method, generalized synthetic controls, that generates counterfactual units based on a latent factor model (Xu 2017).

There is sufficient data to construct control sets for nine treated cities in the fiscal models and six for the city council models in the main models, as Table 1 indicates. In Appendix Section D.2, I run the same models using two-way fixed effects and demographic covariates. This enables me to include the starred cities from Table 1, increasing the treatment groups to fourteen and ten cities respectively. That said, the PanelMatch approach is preferable since it avoids parametric assumptions about the covariates that weaken the interpretation of the two-way fixed effects models.

5 Policy and Ideological Representation

In this section, I explore whether RCV produces governance more representative of the typical voter. If this is the case, we ought to see increased representation on policy outcomes along with a more congruent city council. I start by examining changes in fiscal policy following the adoption of RCV. Since cities vary widely in their fiscal regimes, I log the taxation

[17] I use this method since it produces the largest improvement in covariate imbalance. The findings hold if I construct a matched control set on Mahalanobis distance instead of CBPS weights.

and spending per capita measures (De Benedictis-Kessner and Warshaw 2016). Figure 1 examines the average treatment effects for the full set of municipal revenue and expenditure variables (see Table C.1 for coefficients). In each row, the first coefficient corresponds to changes in policy outcomes. A positive coefficient means that spending or taxes increased following adoption of RCV. Since the dependent variables are estimated in logged dollars per capita, the coefficients should be exponentiated to interpret the effect sizes. For example, a coefficient of 0.15 would suggest that RCV increased spending by $e^{0.15} \approx 1.16$, or 16%. The second coefficient in each row corresponds to changes in representation. For the sake of interpretation, I reverse the sign of the quantity in the analyses such that negative coefficients mean declines in representation (i.e., increases in the distance between actual policy and desired policy). Thus, a negative coefficient means that RCV made the policy more incongruent. A coefficient of -0.15 would mean that the distance between actual policy and preferred policy increased by 16%.

From a theoretical perspective, insofar as policymakers are more extreme than the typical voter and most of the treated cities are left-leaning, we might expect spending and taxation to decrease following adoption of RCV (i.e., negative coefficients on spending variables), but the main theoretical expectation is that RCV improves representation (i.e., positive coefficients on representation variables).

Turning to the actual spending outcomes, RCV does not have a statistically significant effect any of the sixteen variables. In the second set of coefficients, representation declines on police and public welfare spending due to RCV. However, in both sets, we are testing a plethora of dependent variables, which increases the likelihood of false positives. To account for this, I control for the family-wise error rate in each set using a Holm correction. After doing so, only the decline in public welfare spending is statistically significant.

The model suggests that RCV produces a 30% increase in the distance (i.e., a decline in representation) between welfare spending and preferred spending levels. That said, there

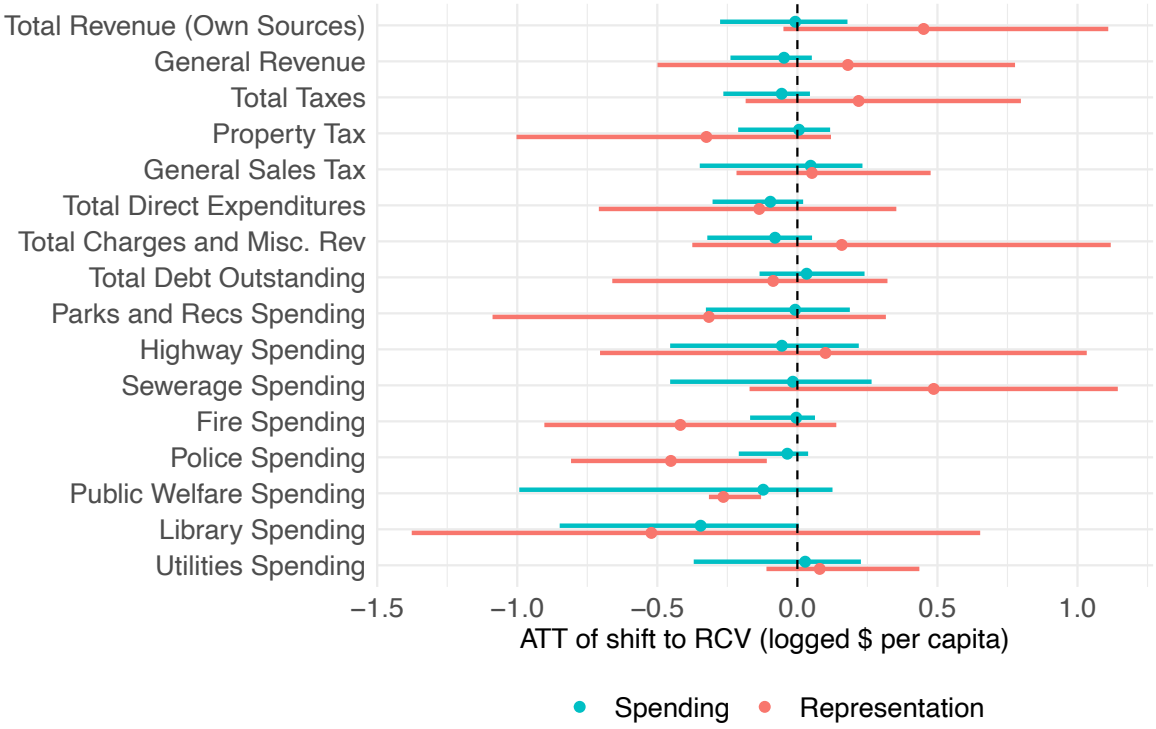

Figure 1: Effects of RCV on levels and representation of taxation and spending. The coefficients indicate the impact of RCV on logged dollars per capita for spending and representation respectively. The bars indicate bootstrapped 95% confidence intervals. In each model, effects are estimated on all nine treated cities except the following due to data missingness: general sales tax (7), sewerage (8), fire (8), public welfare (3), libraries (7), and utilities (6). Representational models do not include Aspen due to a lack of constituency ideology data.

are two reasons to take the finding with a grain of salt. First, there is welfare spending data in only three of the treated cities: San Francisco, Aspen, and Portland. Second, and more importantly, unlike the other fifteen fiscal variables, there is a severe covariate imbalance between the treated and untreated cities. While the weighting procedure reduces the severity of the imbalance, it is still approximately ten times larger here than the imbalance for the other fiscal variables (see Figure B.2). As such, more work should be done to explore whether RCV has negative impacts on welfare policy representation. The primary takeaway from Figure 1 is that RCV has a minimal impact on fiscal outcomes.

That said, it may be the case that policy outcomes do not change because cities face legal and political constraints over their budgets that determine fiscal levels, but that RCV nonetheless induces a shift in the ideological composition of city councils (Peterson 1981). I use the contributor CFscores to test this possibility. These scores mostly take on values from -2 to 2, where higher values are more conservative.[18] I again run two separate models: the first tracks changes in the ideological composition following RCV, and the second explores changes in the distance between the council and mass opinion. Since the city council variables are not logged, the coefficients can be interpreted straightforwardly in the CFscore variable space.

Table 2: Effects of RCV on Ideology and Representation in City Councils

	Dependent variable:	
	Ideology	Representation
RCV Election	-0.018	-0.050
	$(-0.166, 0.216)$	$(-0.187, 0.049)$

Note: Models generated using PanelMatch. For the ideology models, positive scores indicate changes in the conservative direction. 95% bootstrapped confidence shown in parentheses. * indicates $p < 0.05$ (two-tailed tests).

[18] 0.4% of individual legislator scores and 0.01% of city council mean scores are outside of this range.

In Table 2, I show the treatment effect of RCV on city council ideology. The first coefficient indicates a leftward shift of 0.018 units of city council ideology due to RCV, but the effect is clearly not statistically significant and substantively miniscule (0.02 SDs). The negative coefficient in the representational analysis suggests a possible decline in representation, but the coefficient again is not statistically significant and is substantively small (0.06 SDs).

Overall, it does not appear that RCV meaningfully alters the policy outcomes or the ideological composition of city councils. This pattern of null results holds when using generalized synthetic controls and two-way fixed effects models (Appendix Sections D.1 and D.2 respectively). In Appendix Section E, I show that these findings remain when the sample is split between between cities that hold on-cycle versus off-cycle elections.

6 Using Ideal Points from Roll Call Votes

While Figure 1 and Table 2 suggest that policy and ideological representation are unaffected by RCV, two questions persist: one theoretical and one methodological. First, while the findings in Figure 1 tells us that policy representation did not improve under either the first or the second theory, the city council findings in Table 2 cannot distinguish between the two theories because I use static legislator ideal points. In other words, the lack of change could occur due to an unchanged translation of voter preferences or because legislators are not adapting over time (or both are happening in offsetting directions). Second, there may be characteristics of the local politics in these cities that call into question the cross-city comparisons in the representation metric I construct, which assumes a linear relationship between mass liberalism and demand for public goods on any given issue.

To address these questions, I use municipal roll call data. First, I test whether RCV affects representation by *adapting* the behavior of legislators (rather than merely having effects through *replacement*) using dynamically estimated ideal points. If the same legislators

moderate after RCV implementation, the second theory would be supported. Second, I estimate changes in responsiveness *within* city politics following the adoption of RCV. I also explore changes in the council's ideology using roll call-based ideal points as a robustness check to Table 2 in Appendix Section F.2, and I find similar null results.

One limitation is that city council roll call data is not readily available. Bucchianeri (2018) generates dynamic ideal points for the San Francisco Board of Supervisors from 2000 to 2016 to assess the ideological structure of municipal politics in the city, which I reanalyze here. I also estimate dynamic ideal points for Oakland based on roll call data Bucchianeri has collected there (see Appendix F.1 for more detail). As a validity check, the roll-call based ideal point for a legislator (averaging across all of her terms) has a correlation with her CFscore of 0.41 in Oakland and 0.62 in San Francisco, which compares favorably to within-party correlations between ideology metrics at the congressional level (Tausanovitch and Warshaw 2017). Bucchianeri (2018) has estimates of voter ideology at the city district level for San Francisco, and I compare these estimates to the corresponding legislator ideal points to test whether responsiveness changes after RCV adoption.

Table 3: Within-Legislator Shifts in Ideal Points after RCV

	Dependent variable:			
	San Francisco		Oakland	
RCV Election	-0.561^{*}	-0.341	-0.028	0.075
	(0.201)	(0.190)	(0.381)	(0.149)
Year		-0.047		-0.016
		(0.038)		(0.054)
Supervisor Fixed Effects	Yes	Yes	Yes	Yes
Observations	96	96	66	66

Note: Higher values indicate more conservative voting records. Models include standard errors clustered at councilmember level. There are 31 unique councilmembers in San Francisco and 16 in Oakland in the periods examined. * indicates $p < 0.05$ (two-tailed tests).

To evaluate the first claim, I run a regression of RCV adoption on these dynamic ideal

points with legislator fixed effects. Table 3 shows the results of these models (since the ideal points for the two cities are estimated on separate scales, the coefficients are not directly comparable). In the model with RCV alone, the coefficient is positive and significant in San Francisco. However, when the linear time trend is included, the RCV coefficient does not achieve statistical significance. This suggests that RCV did not produce any change beyond the existing trend. Moreover, the positive coefficient points in the *wrong* direction that the theory presupposes; rather than moderating, the positive within-legislator effect suggests that RCV pushed legislators even farther to the left. In Oakland, there is no effect of RCV adoption under either specification, suggesting that legislator behavior did not change following adoption of RCV.

The second question is whether representation within a given city improves following the adoption of RCV. Bucchianeri (2018) uses data on San Francisco municipal public opinion at the city district level based on a factor analysis of ballot proposition returns (DeLeon and Latterman 2004; DeLeon and Latterman 2006; Latterman 2011; Latterman 2015). He finds a positive relationship between district liberalism and supervisor liberalism. I reanalyze this data to determine whether the degree of responsiveness changed following RCV adoption. I regress legislators' ideal points from the Progressive Voter Index (PVI) in the Bucchianeri data, a variable indicating whether the preceding election was held with RCV, and an interaction between the two terms. I flip the direction of the PVI so that higher values map on to more conservative opinion. I look at the statistical significance of the interaction to determine whether responsiveness changed following adoption of RCV (Hagen, Lascher Jr., and Camobreco 2001).

Table 4 shows the results of this model. As in Bucchianeri (2018), I find a positive relationship between district ideology and supervisor ideology in column 1. However, the interaction between PVI and RCV is not significant, suggesting that RCV had no impact on responsiveness. In column 2, I run the same model including a linear time trend and

Table 4: Responsiveness in San Francisco Elections and RCV

	Legislator Ideal Point			
	(1)	(2)	(3)	(4)
Progressive Voter Index	0.023*	0.023*	−0.009	−0.005
	(0.006)	(0.006)	(0.023)	(0.024)
RCV Election	−0.123	−0.241	−0.390	−0.392
	(0.436)	(0.465)	(0.525)	(0.526)
Election Year		0.018		0.008
		(0.019)		(0.019)
Progressive Voter Index × RCV Election	−0.004	−0.004	−0.007	−0.006
	(0.008)	(0.008)	(0.009)	(0.009)
District Fixed Effects	No	No	Yes	Yes
Observations	93	93	93	93

Note: Higher values indicate more conservative voting records. The PVI index is flipped such that higher values indicate more conservative mass opinion. Models include standard errors clustered at district level. There are 11 single-member districts in San Francisco. * indicates $p < 0.05$ (two-tailed tests).

find similar results. Columns 3 and 4 reproduce the models of columns 1 and 2 but include district fixed effects. Here, the effect of PVI on legislator ideal point disappears, suggesting that within-district changes in ideology do not produce changes in municipal representation. The important takeaway here, though, is that the interaction remains insignificant in all models. In sum, it does not appear that RCV affected the caliber of responsiveness in San Francisco.

7 Changes in the Candidate Pool

Although RCV does not alter policy outcomes or city council representation, the third theory suggests that it may alter the candidate pool. I conduct three statistical tests to evaluate this claim: whether the mean, variance, or distribution of the candidate pool shifted following adoption. For the analyses in this section, I use CFscores in the twenty-year window

around adoption for each city. I only include all cities with data for more than ten candidates both before and after adoption.[19]

City	No. of candidates (post, pre)	Diff. in Means	Ratio of Variances	K-S Statistic
Berkeley	(48, 25)	0.05	1.52	0.29
Cary	(11, 20)	-0.27	1.7	0.13
Minneapolis	(53, 37)	-0.15	5.81*	0.10
Oakland	(48, 29)	0.07	3.02*	0.25
San Francisco	(151, 36)	0.1	1.7	0.18
San Leandro	(23, 18)	0.61*	0.33*	0.28
St. Paul	(37, 25)	0.22	0.74	0.26

Table 5: Tests of changes in the candidate distribution. The second column shows the number of candidates in each test post-RCV and pre-RCV. The third column shows the difference in the post-RCV mean and the pre-RCV mean. The fourth column shows the variance of the post-RCV candidate pool over the variance of the pre-RCV pool. The fifth column shows the D-statistic from the Kolmogorov-Smirnov test that the pre- and post-RCV distributions are the same. * indicates $p < 0.05$ (two-tailed tests).

I start by exploring whether the mean of the candidate pool shifted in column 3 of Table 5. I use two-sample t-tests of the mean of the pool before and after adoption. These tests reveal, however, that the mean of the candidate pool shifts in only one of the seven cities, San Leandro (in a conservative direction). This shift barely reaches conventional levels of significance and does not survive a Holm correction. On average, it appears that neither the set of winning legislators nor the candidate pool shift ideologically.

While the means do not appear to have shifted, I move now to a more direct test of advocates' claim that the candidate pool moderates following RCV adoption. To do so, I look at the *variance* of the candidate pool. I use a two-sided F-test of the sample variances in candidate ideology within-city before and after adoption of RCV. Under the null hypothesis, this ratio is 1. If the candidate pool has moderated, we should expect to see lower variance after adoption (and thus a ratio less than 1). If the candidate pool becomes more extreme—

[19]This excludes Portland, ME, which does not have enough data before adoption.

another possible effect of RCV—we would see a ratio greater than 1.

As column 4 shows, the candidate pool variance does not shift significantly for four of these seven observations. The variance increases in Minneapolis and Oakland, but decreases in San Leandro. While the size of these shifts is notable (from a sixfold increase in Minneapolis to a threefold decrease in San Leandro), the results taken as a whole do not point to any systematic shift in ideological variance. To confirm that these findings are not due to year-specific changes coinciding with the adoption of RCV, I verify both the mean and variance results across all city councils (including those that never adopt RCV) using two-way fixed effects models in Appendix Section D.3 and find clearly null effects.

Finally, I consider shifts in the distribution of each city's candidates using a Kolmogorov-Smirnov test in column 5. Rather than testing the mean or variance alone, the K-S test explores whether the entire distribution changes. For every city in the sample, we cannot reject the null hypothesis that the distributions do not change. To summarize, RCV does not have significant impacts on the mean, variance, or distribution of the candidate pools. Across multiple tests, then, it appears that RCV does not even impact the set of candidates who decide to run.

8 Discussion

At this point, it is worth taking stock of the evidence presented relative to the proposed theoretical mechanisms. Recall the three theories by which RCV might improve representation: it might improve the translation of votes holding constant the positions of voters and politicians, it might make politicians and policy outcomes more representative, or it might make the candidate pool more representative.

The findings in Figure 1 and Table 2 show that fiscal and ideological outcomes are largely unchanged following the adoption of RCV. Similarly, the caliber of representation is unaf-

fected; insofar as there was any impact, representation becomes *worse* following adoption. Furthermore, legislative roll-call based responsiveness to city district-level opinion in San Francisco was unchanged following RCV adoption (Table 4), helping to demonstrate that the representation findings in Figure 1 and Table 2 are not driven by the cross-city scaling procedure I use there.

The fiscal results in Figure 1 suggest that both the first and second theories fail on that front. However, we cannot differentiate between the first and second theories in the city council ideology results using Table 2 because we cannot measure shifts in politicians' ideology with static ideal points. In other words, even if we saw the same set of politicians elected post-RCV (a failure on the first theory), it could nonetheless be the case that those politicians are more representative (a success on the second theory) if they adopted more representative platforms. However, Table 3 uses dynamic ideal points from roll-call data to show that individual legislators in San Francisco and Oakland did not change after the adoption of RCV. Thus, it appears that RCV neither improves the translation of voters into winners nor pushes legislators or policy outcomes in a more representative direction.

Finally, it is possible that RCV alters the candidate pool to be more representative even if it does not affect the set of winners—a weaker claim than the first two, but a possibility nonetheless. This claim does not withstand empirical scrutiny, though, as Table 5 shows. The ideological mean of the candidate pools did not shift in these cities after adopting RCV. The candidate pool became neither more ideologically dispersed nor more clustered. Finally, the empirical distribution of candidate ideal points in each city was unchanged following adoption of RCV.

9 Conclusion

Advocates of ranked choice voting point to a host of problems that plague American politics today. They argue that FPTP empowers extremists and deprives voters of meaningful representation. In doing so, these advocates fit into a storied tradition of electoral reformers dating back to the Progressive Era. However, like their Progressive counterparts, it appears that their proposal falls short of meaningful change. In this paper, I provide the first analysis of whether RCV affects substantive representation. I find no evidence that RCV changed the policy outcomes, ideological composition, or responsiveness of city governments.

My results fit into a broader set of findings in the local politics literature. While the core of the representational system holds in local politics—voters are offered distinct choices between candidates who pursue different agendas in office, and voters hold them accountable for the results of their administration—there is little evidence to suggest that the electoral rules structuring municipal politics affect the caliber of this representation. Like term limits, nonpartisan elections, and other reform efforts, RCV is yet another institution that does not live up to the hype.

That said, my findings are limited to the set of largely progressive cities that have adopted the reform. As such, there may be additional moderators that the institutional variation across these cities fails to pick up. These cities may also operate in an ideological and partisan space where electoral reforms will not produce significant changes. It could have an impact in competitive partisan races merely by altering whether a Democrat or Republican wins, as the Maine and Georgia examples demonstrate. As more jurisdictions adopt the reform, others should examine the effects of these reforms, especially where the method differs notably (e.g., Alaska). With more adopters, researchers can also explore heterogeneity in effect sizes. For example, RCV's effects may be concentrated in locations with a certain degree of electoral competition. A related issue for causal identification is that selection into RCV is

non-random. Other work could estimate causal effects based on the plausible exogeneity of treatment (e.g., examining locales which narrowly adopted the reform, court ordered shocks relating to RCV implementation).

While it appears RCV produces limited change in the municipalities that have adopted it, my findings do not resolve several of the outstanding debates between supporters and detractors. For example, enabling voters to cast a sincere ballot may be a normative good in of itself. On the other side, the increased complexity and delays in reporting votes stemming of RCV may make the reform not worth it—especially given that voters prefer FPTP (Nielson 2017). However, at the end of the day, the effects of RCV on municipal policy and representation appear minimal.

References

Abramson, Paul R., John H. Aldrich, André Blais, Matthew Diamond, Abraham Diskin, Indridi H. Indridason, Daniel J. Lee, and Renan Levine (2010). "Comparing strategic voting under FPTP and PR". *Comparative Political Studies* 43 (1): 61–90.

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.

Barber, Michael (2022). "Comparing Campaign Finance and Vote Based Measures of Ideology". *The Journal of Politics* 84 (1): 613–619.

Berry, Christopher R. and William G. Howell (2007). "Accountability and local elections: Rethinking retrospective voting". *Journal of Politics* 69 (3): 844–858.

Bonica, Adam (2014). "Mapping the Ideological Marketplace". *American Journal of Political Science* 58 (2): 367–386.

Bucchianeri, Peter (2018). "There's More than One Way to Party: Progressive Politics and Representation in Nonpartisan San Francisco".

Bucchianeri, Peter (2020). "Party Competition and Coalitional Stability: Evidence from American Local Government". *The American Political Science Review* 114 (4): 1055–1070.

Buisseret, Peter and Carlo Prato (2022). "Politics Transformed? How Ranked Choice Voting Shapes Electoral Strategies".

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.

De Benedictis-Kessner, Justin and Christopher Warshaw (2016). "Mayoral partisanship and municipal fiscal policy". *The Journal of Politics* 78 (4): 1124–1138.

De Benedictis-Kessner, Justin and Christopher Warshaw (2020a). "Accountability for the Local Economy at All Levels of Government in United States Elections". *The American Political Science Review*, 1–17.

De Benedictis-Kessner, Justin and Christopher Warshaw (2020b). "Politics in forgotten governments: the partisan composition of county legislatures and county fiscal policies". *The Journal of Politics* 82 (2): 460–477.

DeLeon, Richard and David Latterman (2004). *Updating the New Progressive Voting Index (PVI), with Tables, Map, and Precinct Scores.*

DeLeon, Richard and David Latterman (2006). *PVI 3.0: Revising & Updating the Progressive Voting Index (PVI) with Maps, Tables, and Precinct Scores.*

Dellis, Arnaud, Alexandre Gauthier-Belzile, and Mandar Oak (2017). "Policy Polarization and Strategic Candidacy in Elections under the Alternative-Vote Rule". *Journal of Institutional and Theoretical Economics* 173 (4): 565.

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

Dynes, Adam M., Michael T. Hartney, and Sam D. Hayes (2021). "Off-Cycle and Off Center: Election Timing and Representation in Municipal Government". *The American Political Science Review* 115 (3): 1097–1103.

Einstein, Katherine Levine and David M. Glick (2018). "Mayors, partisanship, and redistribution: Evidence directly from U.S. mayors". *Urban Affairs Review* 54 (1): 74–106.

Einstein, Katherine Levine and Vladimir Kogan (2016). "Pushing the City Limits: Policy Responsiveness in Municipal Government". *Urban Affairs Review* 52 (1): 3–32.

Ferreira, Fernando and Joseph Gyourko (2009). "Do Political Parties Matter? Evidence from U.S. Cities". *The Quarterly Journal of Economics* (February): 399–422.

Fishburn, Peter C. and Steven J. Brams (1983). "Paradoxes of Preferential Voting". *Mathematics Magazine* 56 (4): 207–214.

Gerber, Elisabeth R. and Daniel J. Hopkins (2011). "When Mayors Matter: Estimating the Impact of Mayoral Partisanship on City Policy". *American Journal of Political Science* 55 (2): 326–339.

Hagen, Michael G., Edward L. Lascher Jr., and John F. Camobreco (2001). "Response to Matsusaka: Estimating the effect of ballot initiatives on policy responsiveness". *The Journal of Politics* 63 (4): 1257–1263.

Hall, Andrew B. (2019). *Who Wants to Run?* Chicago, IL: University of Chicago Press.

Hall, Andrew B. and James M. Snyder (2015). "Information and wasted votes: A study of U.S. primary elections". *Quarterly Journal of Political Science* 10 (4): 433–459.

Hopkins, Daniel J. and Lindsay M. Pettingill (2018). "Retrospective Voting in Big-City US Mayoral Elections". *Political Science Research and Methods* 6 (4): 697–714.

Imai, Kosuke, In Song Kim, and Erik H. Wang (2021). "Matching Methods for Causal Inference with Time-Series Cross-Sectional Data". *American Journal of Political Science* 0 (0): 1–19.

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.

Kilborn, Mitchell and Arjun Vishwanath (2022). "Public Money Talks Too: How Public Campaign Financing Degrades Representation". *American Journal of Political Science* 66 (3): 730–744.

Kilgour, D. Marc, Jean Charles Grégoire, and Angèle M. Foley (2020). "The prevalence and consequences of ballot truncation in ranked-choice elections". *Public Choice* 184 (1-2): 197–218.

Kimball, David C. and Joseph Anthony (2016). "Voter Participation with Ranked Choice Voting in the United States".

Kousser, Thad, Justin Phillips, and Boris Shor (2018). "Reform and Representation: A New Method Applied to Recent Electoral Changes". *Political Science Research and Methods* 6 (4): 809–827.

Latterman, David (2011). *PVI 5.0: Revised 2011 San Francisco Progressive Voter Index and Changes in the San Francisco Electorate*.

Latterman, David (2015). *PVI 6.0: Revised 2015 San Francisco Progressive Voter Index and Changes in the San Francisco Electorate*.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles (2021). *IPUMS National Historical Geographic Information System: Version 16.0 [dataset]*. Minneapolis, MN.

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.

Nielson, Lindsay (2017). "Ranked Choice Voting and Attitudes toward Democracy in the United States: Results from a Survey Experiment". *Politics and Policy* 45 (4): 535–570.

Ornstein, Joseph T. and Robert Z. Norman (2014). "Frequency of monotonicity failure under Instant Runoff Voting: Estimates based on a spatial model of elections". *Public Choice* 161 (1-2): 1–9.

Peterson, Paul (1981). *City Limits*. Chicago, IL: University of Chicago Press.

Pierson, Kawika, Michael L. Hand, and Fred Thompson (2015). *The Government Finance Database: A Common Resource for Quantitative Research in Public Financial Analysis*.

Richie, Robert (2004). "Instant Runoff Voting: What Mexico (and Others) Could Learn". *Election Law Journal* 3 (3): 501–512.

Robb, Denise Munro (2011). "The Effect of Instant Runoff Voting on Democracy". PhD thesis. University of California, Irvine.

Rogers, Steven (2017). "Electoral Accountability for State Legislative Roll Calls and Ideological Representation". *The American Political Science Review* 111 (3): 555–571.

Santucci, Jack (2018). "Evidence of a winning-cohesion tradeoff under multi-winner ranked-choice voting". *Electoral Studies* 52 (February): 128–138.

Santucci, Jack and Jamil Scott (2021). "Do Ranked Ballots Stimulate Candidate Entry?"

Shor, Boris and Nolan McCarty (2011). "The Ideological Mapping of American Legislatures". *The American Political Science Review* 105 (3): 530–551.

Tausanovitch, Chris and Christopher Warshaw (2014). "Representation in Municipal Government". *The American Political Science Review* 108 (3): 605–641.

Tausanovitch, Chris and Christopher Warshaw (2017). "Estimating candidates' political orientation in a polarized congress". *Political Analysis* 25 (2): 167–187.

Warshaw, Christopher, Justin De Benedictis-Kessner, Daniel M. Thompson, and Yamil Velez (2022). "Local Representation in the United States: A New Comprehensive Dataset of Elections".

Xu, Yiqing (2017). "Generalized synthetic control method: Causal inference with interactive fixed effects models". *Political Analysis* 25 (1): 57–76.

ONLINE APPENDIX

Robustness Checks and Supplementary Analyses for

Electoral Institutions and Substantive Representation in Local Politics: The Effects of Ranked Choice Voting

Contents

A Evaluating Translation of Voter Preferences with Simulations **1**

B Covariate Balance Tests **5**

C Regression Coefficients from PanelMatch models **7**

D Robustness Checks **9**

D.1 Generalized Synthetic Control Models 9

D.2 Two-Way Fixed Effects Models 12

D.3 Alternative Approach to Candidate Pool Moderation 14

E Heterogenous Effects by Election Timing **15**

F Roll Call Based Ideal Points **18**

F.1 Details of Ideal Point Estimation in Oakland 18

F.2 Changes in San Francisco and Oakland City Council Ideology Following RCV Adoption 19

A Evaluating Translation of Voter Preferences with Simulations

In this section, I examine the hypothesis that RCV improves the translation from voters' preferences to winners. To test the claim, I run simulations of various electorates and candidate sets with randomly generated ideal points. The logic of spatial models suggests that voters' and candidates' positions can be represented in Euclidean space. In these models, voters prefer candidates whose ideal points are closer to their own. The goal of my analysis is to determine which system produces the outcome favored by the median voter.

To do so, I employ a one-dimensional spatial voting model under three electoral systems. Under the first, simple FPTP, each voter votes for the candidate closest to her ideal point, and the candidate with the most votes wins. Thus, if the voter's ideal point is 0.1 and the candidates in the race have ideal points of $A = 0.4$, $B = -0.8$, and $C = 0.3$, she will vote for C. While this helps to examine the plurality winner problem, it does not deal with the question of sincere versus strategic voting. To address this, I analyze a second system ("restricted FPTP") in which voters only consider casting a ballot for the first or second candidate, regardless of the other candidates' ideal points. This rule mimics the national status quo, where voters expect one of the two major party candidates to win almost every election and so they disregard the other choices available to them.[1] Under this rule, our hypothetical voter would vote for A since C is non-viable and she prefers A to B. In the third system, voters rank a full ballot and the winner is chosen via RCV. In our example, the voter ranks $C \succ A \succ B$.

For these three election rules, I run models with anywhere between 3 and 8 candidates and two different 1,001-voter electorates. The first electorate is a unimodal distribution and the other is a "polarized" bimodal distribution.[2] As an example, Figure A.1 shows the results of two separate six candidate simulations, one of each electorate type. In each panel, the density plot shows the distribution of voters (the top panel is the unimodal electorate and the bottom panel is the polarized electorate). The blue vertical lines represent the candidates' positions in a six-candidate field. Below each density plot, the label "Med. Pref" indicates the preferred candidate of the median voter, and the labels "FPTP", "FPTP Restr.", and

[1] Since the candidates' ideal points are randomly drawn from the same distribution, the choice to use the first and second candidates generated is made without loss of generality. This assumption will lead me to underestimate the representativeness of FPTP because the major parties do exhibit some nonzero degree of responsiveness in their platforms to public opinion. I make the simplifying assumption that all voters select one of the major party candidates, but my findings remain if a small number of voters cast sincere ballots for other candidates.

[2] For the former, I generate 1,001 draws from an $N(0.5, 0.2^2)$ distribution. For the latter, I generate two means, $\mu_l \sim \text{Unif}(0.1, 0.3)$ and $\mu_r \sim \text{Unif}(0.7, 0.9)$. Then, I combine 500 draws from a $N(\mu_l, 0.1^2)$ distribution with 501 from a $N(\mu_r, 0.1)$ distribution. By first simulating the means, I let the bimodal distributions be asymmetric.

"RCV" indicate the winners under each of those three systems.

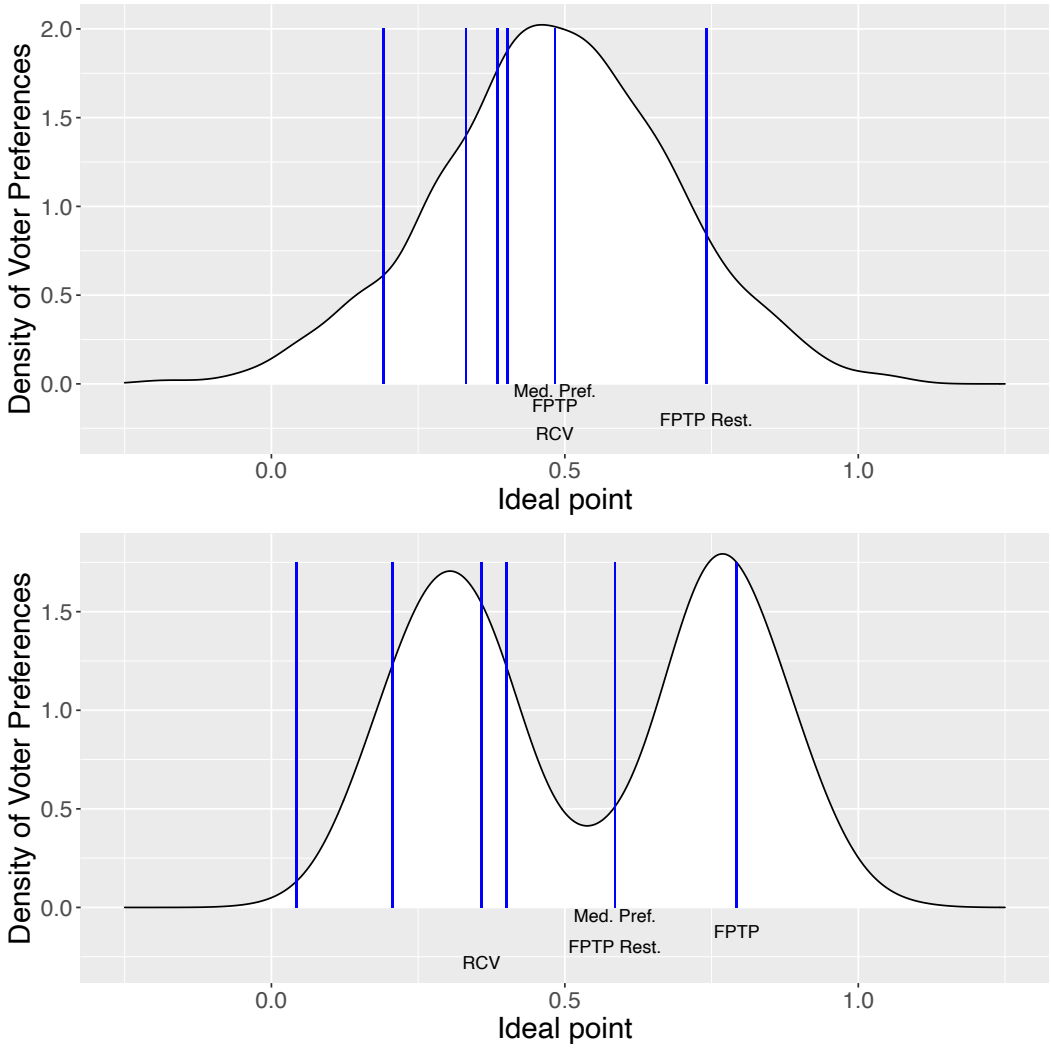

Figure A.1: Two simulations under a unimodal electorate (top panel) and a polarized electorate (bottom panel). In the unimodal example, the median voter prefers the candidate with ideal point 0.48. RCV selects this candidate (as does simple FPTP), but restricted FPTP selects the candidate with ideal point 0.74. In the polarized example, the median voter prefers the candidate with ideal point 0.59. Restricted FPTP selects this candidate, while RCV selects a winner with ideal point 0.36 (simple FPTP selects one with ideal point 0.79).

For each number of candidates and electorate type, I run 1,000 simulations and record whether each of the three election rules produces the winner preferred by the median voter. Figure A.2 shows the probability that each rule does so. Starting with the unimodal electorate in the left panel, we see that RCV produces a median-preferred winner more often

than FPTP or restricted FPTP does.[3] However, when we turn to the polarized electorate, RCV's advantage disappears. Restricted FPTP always performs better than RCV. In fact, with six or more candidates, RCV is less likely to produce the median's preferred candidate than *if the winner were randomly drawn from a hat.*

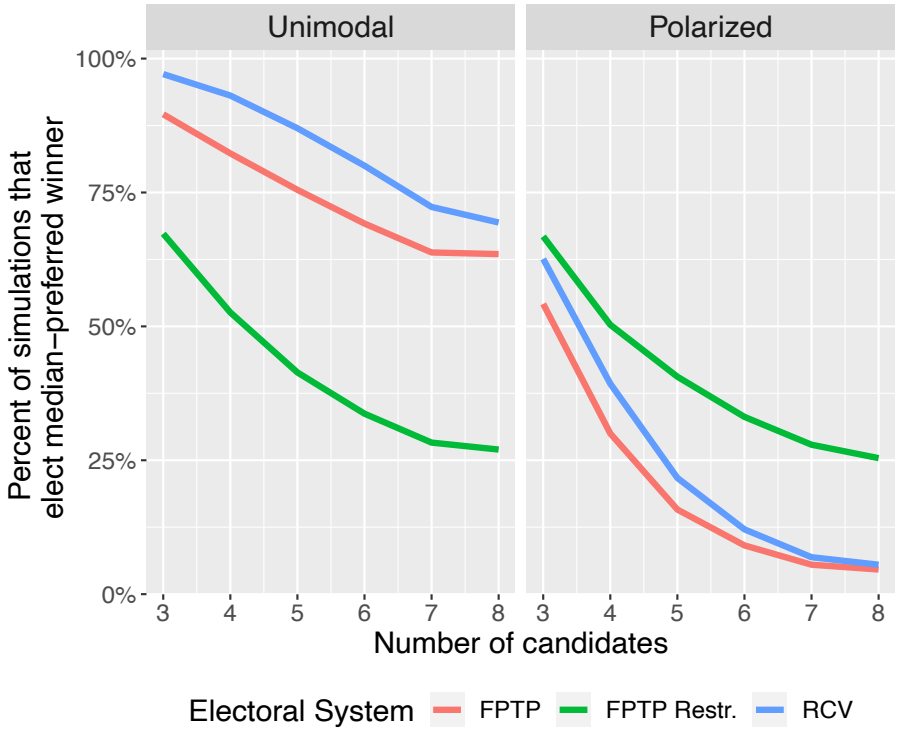

Figure A.2: Success rates by electorate type and number of candidates

Why is this the case? In these elections, RCV often eliminates the median-preferred candidate in early rounds.[4] Of course, these simulations are sensitive to the electorate

[3]When the median-preferred candidate is Candidate A or Candidate B (i.e., one of the two major party candidates), restricted FPTP almost always identifies the median-preferred candidate, but in every other case, it necessarily fails. As a result, restricted FPTP's success rate is approximately equal to the probability that one of the restricted candidates is the median's preference, or $\frac{2}{N_{cand}}$.

[4]To see why this is more likely in a polarized electorate, consider an 11 person electorate with candidates L (left-wing), C (centrist), and R (right-wing). Five voters have the preference profile $L \succ C \succ R$, another five have the profile $R \succ C \succ L$, and the last voter has the profile $C \succ L \succ R$. C is eliminated in the first round, and the last voter's ballot is reallocating to L, shifting the election to L. However, C was the most representative in this election. In a restricted electorate, if C is one of the restricted two candidates, C will win. As the number of candidates increases, it becomes increasingly likely that the moderate

specifications—changing the distributions of voters will produce different findings. Perhaps most importantly, it assumes a constant candidate pool. If RCV moderates the candidate pool, then the translation of voter preferences may be improved by that mechanism. While this exercise analyzes the translation of preferences holding these preferences constant, the decision is sensible given the empirical findings in Table 5 that the candidate pool does not change post-RCV. In sum, we have no theoretical reason to believe that RCV is necessarily more or less likely to translate voter preferences into median-preferred outcomes.

candidates get squeezed out in the early rounds.

B Covariate Balance Tests

Covariate	Ever-treated Mean	Never-treated Mean	Diff.	St. Err of Diff.	p value
Total Population	159,618	9,557	150,060	10,282	<0.01
Percent Male	49.0%	48.8%	0.2%	0.1%	0.14
Percent White	62.0%	81.9%	-19.9%	1.1%	<0.01
Percent Black	9.9%	7.4%	2.5%	0.5%	<0.01
Percent Asian	9.2%	1.2%	8.0%	0.5%	<0.01
Percent Hispanic	15.6%	6.9%	8.7%	0.8%	<0.01
Percent College Plus	46.2%	18.4%	27.8%	0.6%	<0.01
Median Income	$54,717	$44,469	$10,247	$853	<0.01
Median Home Value	$364,553	$123,883	$240,669	$12,246	<0.01
Proportion Homeowners	53.0%	72.7%	-19.7%	0.5%	<0.01
Democratic Vote Share (County)	67.5%	42.6%	24.8%	0.6%	<0.01

Table B.1: Raw covariate balance between cities that ever adopted RCV and those that never did. The second and third columns show means for cities that ever adopted RCV and those that never did. The fourth column shows the difference between the second and third columns. The fifth column shows the standard error of the difference, while the sixth column shows whether the difference is statistically significant. Observations are at the city-year level.

Table B.1 shows the raw differences between the cities that ever adopted RCV and those that never did. On every variable other than percent male, there are statistically significant differences between the treated cities and untreated cities.

Next, Figure B.1 shows the improvement in balance before and after implementing covariate balancing propensity score weights for the spending and ideology variables. Figure B.2 shows the same for the representation estimates of those variables. I take the absolute value of the standardized difference across each of the covariates shown in Table B.1 between the treatment and control groups for a given dependent variable. The estimates differ by dependent variable insofar as data does or does not exist for a given dependent variable. For example, many cities do not report public welfare spending, and so the comparison of the covariates is between control and treatment cities that have reported public welfare spending in the relevant row. Nonetheless, the figure consistently shows a significant improvement in the weighted dataset. Due to data limitations, the public welfare spending variable still sees a significant amount of imbalance—especially on the representational side—which I discuss in the main text.

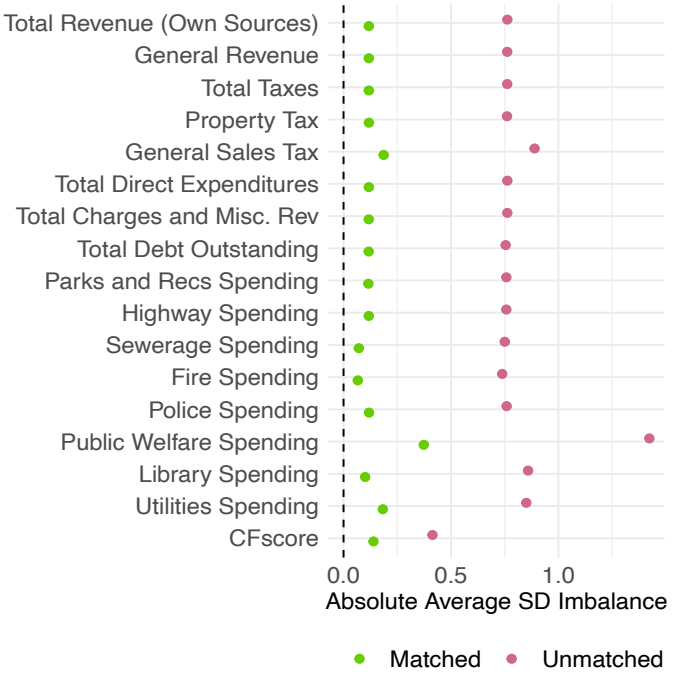

Figure B.1: Covariate balance before and after covariate balancing propensity score weights, spending and ideology variables

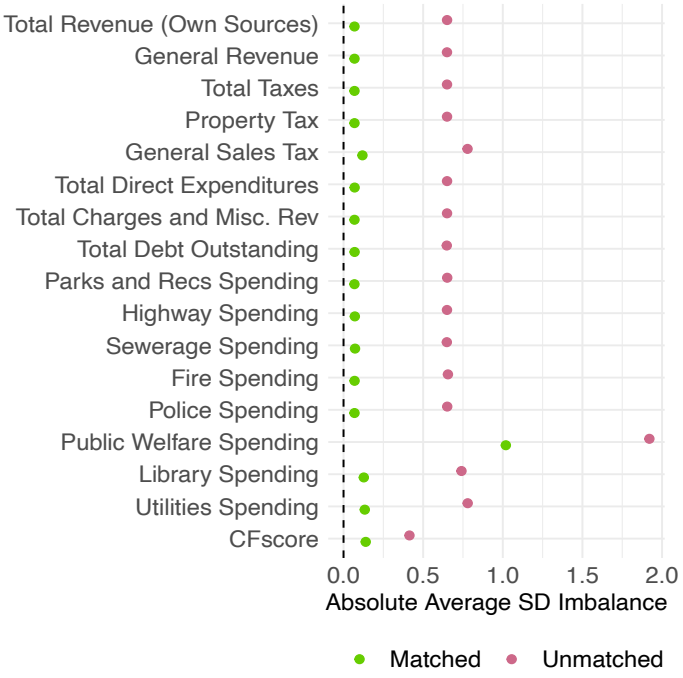

Figure B.2: Covariate balance before and after covariate balancing propensity score weights, representation variables

C Regression Coefficients from PanelMatch models

In this section, I present the coefficients from the main analyses in the generalized synthetic controls. Tables C.1 and C.2 contain the coefficients mapping to spending and representation respectively in Figure 1.

Dependent Variable	ATT	95% CI
Total Revenue (Own Sources)	-0.01	(-0.28, 0.18)
General Revenue	-0.05	(-0.24, 0.05)
Total Taxes	-0.06	(-0.26, 0.04)
Property Tax	0.01	(-0.21, 0.12)
General Sales Tax	0.05	(-0.35, 0.23)
Total Direct Expenditures	-0.10	(-0.3, 0.02)
Total Charges and Misc. Rev	-0.08	(-0.32, 0.05)
Total Debt Outstanding	0.03	(-0.14, 0.24)
Parks and Recs Spending	-0.01	(-0.33, 0.19)
Highway Spending	-0.06	(-0.45, 0.22)
Sewerage Spending	-0.02	(-0.45, 0.26)
Fire Spending	-0.00	(-0.17, 0.06)
Police Spending	-0.04	(-0.21, 0.04)
Public Welfare Spending	-0.12	(-0.99, 0.13)
Library Spending	-0.35	(-0.85, 0)
Utilities Spending	0.03	(-0.37, 0.23)

Table C.1: Coefficients from fiscal spending models

Dependent Variable	ATT	95% CI
Total Revenue (Own Sources)	0.45	(-0.05, 1.11)
General Revenue	0.18	(-0.5, 0.78)
Total Taxes	0.22	(-0.18, 0.8)
Property Tax	-0.32	(-1, 0.12)
General Sales Tax	0.05	(-0.22, 0.48)
Total Direct Expenditures	-0.14	(-0.71, 0.35)
Total Charges and Misc. Rev	0.16	(-0.38, 1.12)
Total Debt Outstanding	-0.09	(-0.66, 0.32)
Parks and Recs Spending	-0.32	(-1.09, 0.32)
Highway Spending	0.10	(-0.7, 1.03)
Sewerage Spending	0.49	(-0.17, 1.14)
Fire Spending	-0.42	(-0.9, 0.14)
Police Spending	-0.45	(-0.81, -0.11)
Public Welfare Spending	-0.26	(-0.32, -0.13)
Library Spending	-0.52	(-1.38, 0.65)
Utilities Spending	0.08	(-0.11, 0.44)

Table C.2: Coefficients from fiscal representation models

D Robustness Checks

D.1 Generalized Synthetic Control Models

I employ two alternative methods to the PanelMatch models. The first uses generalized synthetic controls (Xu 2017). Unlike traditional synthetic controls, this method permits multiple units treated at different points in time. It explicitly considers heterogenous treatment effects, which is especially useful since there is variation in implementation (e.g., ballot truncation rules) and in institutional contexts (e.g., safe Democratic areas versus more competitive areas). It does so by estimating a latent factor space to construct counterfactual estimates in the pre-treatment period, and then it uses these latent factors to estimate counterfactuals in the post-treatment period for each unit. Since the model employs a staggered adoption framework, I drop all post-repeal observations for the cities that repeal RCV. For estimation purposes, I remove all control cities with population under 1,000. I permit each model to include between 0 and 2 latent factors. The model considers twelve treatment cities for the fiscal models (Aspen, Burlington, Cary, Hendersonville, Berkeley, Minneapolis, Oakland, Portland, San Francisco, San Leandro, Santa Fe, and St. Paul) and five treatment cities for the ideology models (Berkeley, Minneapolis, Oakland, San Leandro, and St. Paul).

Figure D.1 shows the coefficient estimates. Here, it appears that library spending declined. On the representation analyses, representation on property taxes and police spending appears to have worsened. After Holm corrections, though, the library spending decline does not remain significant. The decline in representation on property taxes does remains statistically significant, though.

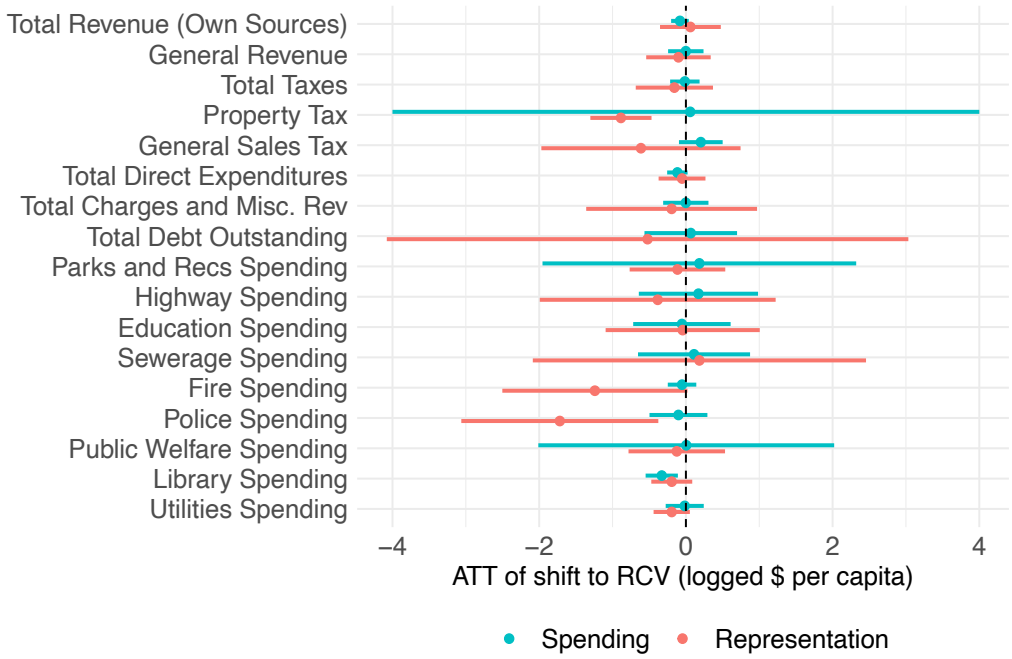

Figure D.1: Effects of RCV on levels and representation of taxation and spending from generalized synthetic control model. The coefficients indicate the impact of RCV on logged dollars per capita for spending and representation respectively. The bars indicate 95% confidence intervals. In each model, effects are estimated on all twelve treated cities except the following due to data missingness: general sales tax (8), highways (11), education (2), public welfare (2), libraries (9), and utilities (9). Representational models do not include Aspen or Hendersonville due to a lack of constituency ideology data. The average number of control cities in each model is 2470. Due to non-convergence in the bootstrapped sample for property taxes, I limit the confidence interval to [-4, 4].

Table D.1: Effects of RCV on Ideology and Representation in City Councils

	Dependent variable:	
	Ideology	Representation
RCV Election	−0.004	−0.035
	(−0.209, 0.201)	(−0.183, 0.112)

Note: Models generated using generalized synthetic controls. For the ideology models, positive scores indicate changes in the conservative direction. 95% confidence shown in parentheses. * indicates $p < 0.05$ (two-tailed tests).

In Table D.1, we can see that city council means did not move nor did the caliber of representation change. In sum, it does not appear that RCV induced any clear changes to representation even under an alternative methodology.

D.2 Two-Way Fixed Effects Models

The second method employs a two-way fixed effects setup with the same covariates as a vector of controls. In other words, rather than weighting, I use the entire set of untreated cities as the control group and control for relevant demographic and political differences. A downside of this approach is that two-way fixed effects are not well suited to imbalanced datasets in which the treated units vary substantially from the control units, which is the case here as Appendix Section B shows. While city fixed effects control for part of this variation, they cannot account for the question of whether the treatment and control samples come from the same population. Another issue is that these city fixed effects may not account for shifts that take place within city over the quarter century studied. However, a key benefit of running the two-way fixed effects model is that it permits inclusion of cities with fewer lag and lead periods. With this analysis, there are fourteen cities included in the fiscal analyses and ten in the city council analyses (see starred entries in Table 1). Four of the fourteen cities are dropped from the fiscal representational analyses due to missingness of public opinion data.

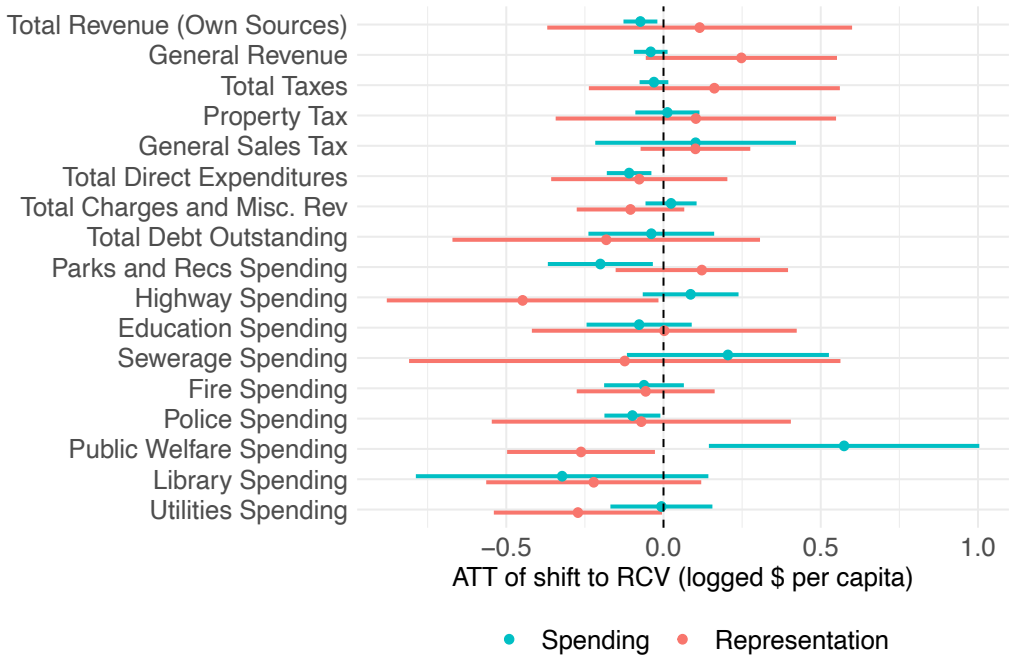

Figure D.2: Effects of RCV on levels and representation of taxation and spending from two-way fixed effects model. In each spending model, effects are estimated on all 14 treated cities except the following due to data missingness: general sales tax (12), education (5), fire (12), public welfare (6), and libraries (9). In each representation model, effects are estimated on all 10 treated cities except for the following: general sales tax (9), education (4), public welfare (6), and libraries spending (9). The average number of control cities in each model is 8,206.

Figure D.2 shows the treatment effects on the spending outcomes using two-way fixed effects. In this model, five of the coefficients are statistically significant: total revenue, direct expenditures, parks and recreation spending, and police spending declined while public welfare spending increased. However, after a Holm correction due to multiple hypothesis testing, only the decline in direct expenditures is significant. Similarly, there are representational shifts on highway spending, public welfare spending, and utilities spending, but none of these effects survive the Holm correction.

Table D.2: Effects of RCV on Ideology and Representation in City Councils, Two-Way Fixed Effects

	Dependent variable:	
	Ideology	Representation
RCV Election	−0.020	−0.039
	(0.052)	(0.045)

Note: For the ideology models, positive scores indicate changes in the conservative direction. Fixed effects by year and city included. Controls included for income, gender, race, education, home value, homeownership, and county presidential vote share. * indicates $p < 0.05$ (two-tailed tests).

Table D.2 shows the treatment effects for the city council ideology outcomes. Neither of the the coefficients is statistically significant. It appears that generally speaking, the main findings are robust to the inclusion of more treated cities.

D.3 Alternative Approach to Candidate Pool Moderation

Here, I verify the findings from the difference-in-means and F tests in Table 5. It is possible that we cannot simply compare the pools of candidates before and after RCV is implemented because of time-varying characteristics across all cities. To account for this, I employ a two-way fixed effects analysis with both the treated and control cities to see whether the sample variance of the candidate pool decreases following adoption of RCV. If this is the case, then we would have reason to believe that the reform had a moderating effect. If the variance were to increase, that would suggest that candidates became more polarized after the adoption of RCV. In the analysis, each observation is at the city council-year level.

Table D.3: Changes in candidate composition post-RCV

	Dependent variable:		
	Variance (Linear)	Variance (Gamma)	Mean
RCV Election	−0.044	0.071	−0.038
	(0.062)	(0.222)	(0.082)
Observations	3,640	3,637	3,640

Note: All models have city and year fixed effects. Models include standard errors clustered at city level. Mean models restricted to city-years with at least 2 observations, gamma models restricted to cities with non-zero variance. * indicates $p < 0.05$ (two-tailed tests).

Table D.3 shows the results of this test. Each observation is aggregated at the city-year level. The first column tests the variance of the CFscore pool. It does not appear that the variance decreased following the adoption of RCV. That said, since the sample variance (the dependent variable here) is non-normally distributed, the use of linear regression may not be appropriate. If the underlying data is (iid) normally distributed, the sample variance takes on a χ^2 distribution, which is a special case of the Gamma distribution. I rerun the model using a Gamma regression in columns 2, and once again, I do not find statistical significance. Furthermore, I verify my findings that the mean of the candidate pool did not change post-RCV in column 3.

E Heterogenous Effects by Election Timing

One possibility is that there are systematic differences in off-cycle versus on-cycle cities, as Dynes, Hartney, and Hayes (2021) suggest. They find that on-cycle cities are more responsive (cross-sectionally) to public opinion than off-cycle cities are. Thus, the null findings in the manuscript might be attributed to heterogeneity between on-cycle and off-cycle cities. If this is the case, we might expect representational gains to be concentrated in off-cycle cities relative to on-cycle cities. There are two reasons this may take place. First, since off-cycle cities start from a lower baseline level of responsiveness, they have more room to improve. Second, it is easier for concentrated interest groups to mobilize in off-cycle elections in favor of their own interests versus the diffuse majority's (Anzia 2011). RCV may force campaigns to appeal more broadly beyond these interests.

I use the coding of cities from Dynes, Hartney, and Hayes (2021) into on-cycle and off-cycle to split the sample into two groups. Of the cities used in the GSC models, 1,182 held off-cycle elections, 329 held on-cycle elections, and there was no data for 18,119 of them. Of the RCV adopting cities, Berkeley, Oakland, San Francisco, and San Leandro hold elections on-cycle, while Portland, Minneapolis, St. Paul, Cary, and Burlington hold elections off-cycle (there is insufficient fiscal data for San Leandro or city council data for Burlington). Because there is only one city with public welfare spending data in the off-cycle and on-cycle sets respectively, I omit it from these analyses.

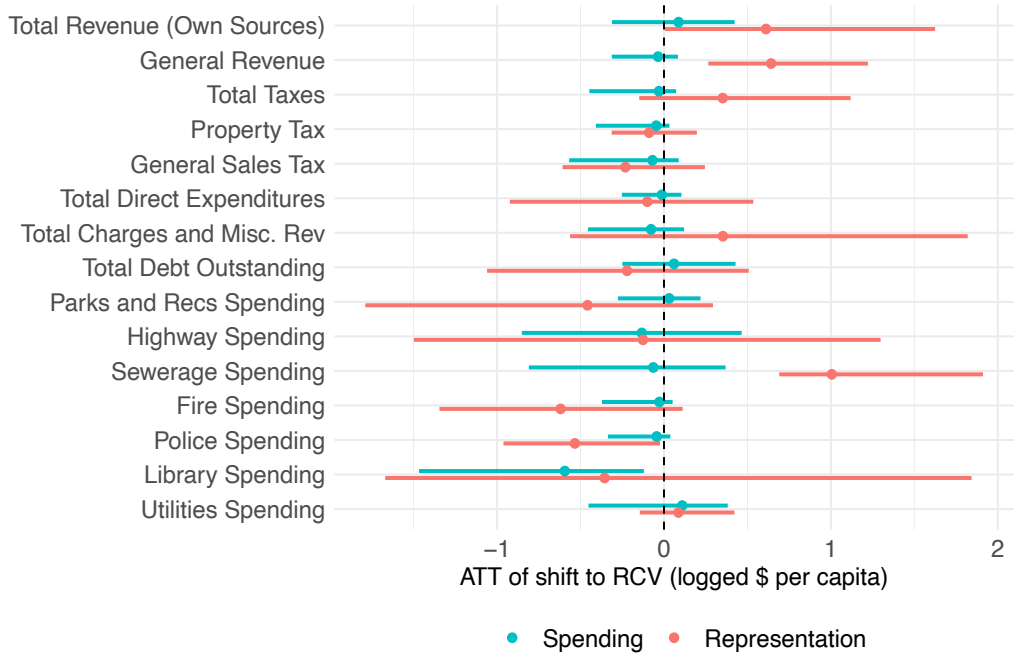

Figure E.1: Effects of RCV on levels and representation of taxation and spending (off-cycle election cities). In each model, effects are estimated on all five treated cities except the following due to data missingness: general sales tax (3), libraries (4), and utilities (4).

Figure E.2: Effects of RCV on levels and representation of taxation and spending (on-cycle election cities). In each model, effects are estimated on all three treated cities. Models could not be estimated for utilities spending due to insufficient data on treated units.

Figure E.1 shows the analog of Figure 1 with the data subset to cities holding elections off-cycle. There is a statistically significant decrease in library spending, but it does not persist after a Holm correction. There are also representational gains in total revenue, general revenue, and sewerage spending along with representational declines in police spending. After a Holm correction, only the gains in general revenue and sewerage spending remain significant. Figure E.2 shows the equivalent for on-cycle cities. There is an increase in parks and recreation spending, but it loses significance after a Holm correction. There is also a representational decline on library spending and parks and recreation spending, but only the library spending coefficient retains significance after a Holm correction.

Table E.1 shows the findings for the ideological variables. There appear to be minimal changes in the off-cycle cities in terms of ideology or representation. Turning to the on-cycle cities, there is a slight ideological shift to the left, which produces a decline in representation that just achieves statistical significance ($p = 0.05$ exactly). However, even this decline is substantively small (0.14 SDs). Insofar as there is any pattern, it seems that representation might have improved in off-cycle cities relative to on-cycle cities, but these differences do not appear meaningful. That said, given the small sample sizes in this analysis, more work should be done to verify the findings.

Table E.1: Effects of RCV on Ideology and Representation in City Councils

	Dependent variable:			
	Off-Cycle		On-Cycle	
	Ideology	Representation	Ideology	Representation
RCV Election	0.028	−0.010	−0.040	−0.117*
	(−0.132, 0.301)	(−0.271, 0.085)	(−0.225, 0.165)	(−0.268, −0.001)

Note: Models generated using PanelMatch. For the ideology models, positive scores indicate changes in the conservative direction. 95% bootstrapped confidence shown in parentheses. * indicates $p < 0.05$ (two-tailed tests).

F Roll Call Based Ideal Points

F.1 Details of Ideal Point Estimation in Oakland

Although Bucchianeri (2018) estimates dynamic ideal points for San Francisco, which I use directly, he does not do so for Oakland. To estimate these for Oakland, I employ roll call data from Oakland's city council that Bucchianeri has collected. As starting values, I let Libby Schaaf take a value of -1 and Rebecca Kaplan take a value of 1. I make these choices based on the endorsement patterns of various legislative candidates. Schaaf ran as a mainstream liberal and received the endorsement of the Oakland Chamber of Commerce along with mainstream Democrats like Jerry Brown.[5] By comparison, Kaplan ran on an explicitly progressive platform and received endorsements from left wing groups like Oakland Rising Action along with national figures like Bernie Sanders.[6] I follow Bucchianeri in running an MCMC model with 10,000 simulations (burn-in of 4,000 simulations) using the `MCMCdynamicIRT1d` function in the `MCMCpack` package. As mentioned in the main text, the average within-legislator score (across all of her terms) correlates with her CFscore at 0.41.

[5] http://www.smartvoter.org/2010/11/02/ca/alm/vote/schaaf_l/endorse.html

[6] http://kaplanforoakland.org/endorsements/

F.2 Changes in San Francisco and Oakland City Council Ideology Following RCV Adoption

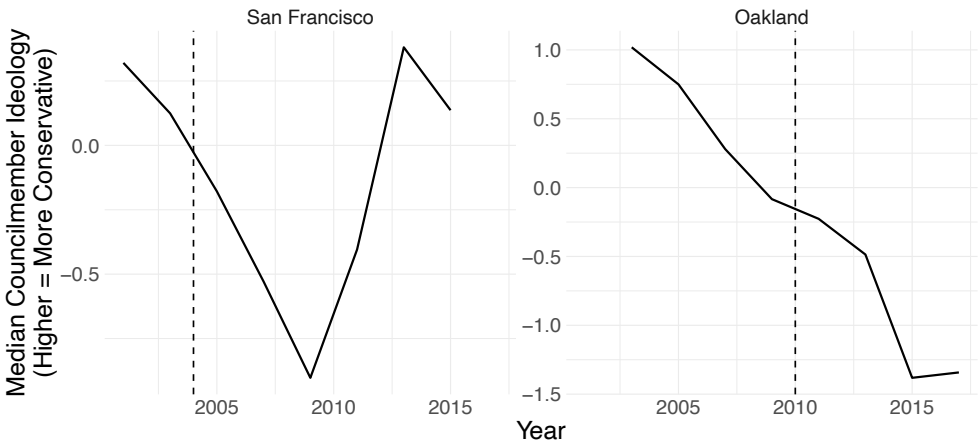

Figure F.1: City council median over time, where higher scores indicate more conservative ideology. The dashed line indicates adoption of RCV. Data points shown are for first year of legislative term (e.g., data from 2005-2006 session is recorded at $t = 2005$). Note that ideal points for each city are estimated separately (since there are no items to jointly scale the two cities), and so the ideal point values should not be compared across cities but only within cities across time.

I explore the shift in the council median before and after RCV as a robustness check to the findings in Table 2. Figure F.1 shows the change in the chamber median over time for each city council, where the vertical dashed line indicates adoption of RCV. In San Francisco, the Board of Supervisors becomes notably more progressive from 2000 to 2010, although as Bucchianeri (2018) notes, this is primarily due to to shifting dynamics in local politics in the period with the rise of the city's progressive movement. In Oakland, the city council becomes consistently more liberal over time as well. Though we cannot decisively determine the casual impact of RCV from Figure F.1 alone, the lack of changes in slope post-treatment suggests that RCV's effect is minimal.[7] The roll-call based estimates, in other words, are consistent with the more robustly estimated findings of Table 2.

[7]To test this more formally, I construct a linear regression in each city predicting the median as a function of time with an interaction term for treatment status. In both models, the interaction is never statistically significant. I also test each model using different post-treatment period lengths (e.g., estimating until 2009 for San Francisco) and the results are unchanged. Of course, the sample size in the regression is small, so these results should be interpreted with caution.