Estimating Candidate Support in Voting Rights Act Cases: Comparing Iterative EI & EI-RxC Methods (*Forthcoming, Sociological Methods & Research*)

Matt Barreto, Loren Collingwood*, Sergio Garcia-Rios, and Kassra Oskooii

February 27, 2019

*Corresponding Author: loren.collingwood@ucr.edu, author order alphabetical

Abstract

Scholars and legal practitioners of voting rights are concerned with estimating individual-level voting behavior from aggregate-level data. The most commonly used technique, King's Ecological Inference (EI), has been criticized for inflexibility in multiethnic settings, or with multiple candidates. One method for estimating vote support for multiple candidates in the same election is called ecological inference: row by columns (RxC). While some simulations suggest that RxC may produce more accurate estimates than the iterative EI technique, there has not been a comprehensive side-by-side comparison of the two methods using real election data that analysts and legal practitioners often rely upon in courts. We fill this void by comparing iterative EI and RxC models in a variety of RxC combinations including two candidates and two groups, three candidates and three groups, and up to twelve candidates and three groups, and multiple candidates and four groups. Additionally, we examine the two methods with 500 simulated datasets that differ in combinations of heterogeneity, polarization, and correlation. Finally, we introduce a new Model Congruence Score (MCS) to further aid the substantive interpretation of the estimates. Across all of our analyses, we find that both methods produce substantively similar results. This suggests that iterative EI and RxC can be used interchangeably when assessing precinct level voting patterns in Voting Rights Act cases, and that neither method produces bias in favor or against finding racially polarized voting patterns.

Introduction

American politics scholars and the U.S. court system commonly assess whether racially polarized voting (RPV) exists in a particular jurisdiction—whether a legislative district, city district, or county supervisor seat—as part of a voting rights analysis. V.O. Key's seminal study of Southern politics documented that Anglos (whites) living around high percentages of Blacks voted most consistently for racially hostile Anglo candidates (Key, 1949). Since then, extensive research has demonstrated that African Americans, Latinos, and Anglos disproportionately favor co-ethnic candidates and exhibit different preferences and voting patterns (Barreto, 2007, 2010; Barreto et al., 2005; Dawson, 2003; Grofman, 1991; Grofman and Handley, 1989; Grofman and Migalski, 1988; Issacharoff, 1992; McCrary, 1990; Piston, 2010; Tate, 1994). With the passage of the 1965 Voting Rights Act (VRA) and subsequent amendments and court decisions, systematic examination of RPV patterns not only became increasingly relevant to scholars of race and ethnic politics, but also the courts and legal practitioners as one major goal of the law was to increase African American voter registration and representation (Cox and Miles, 2008; Davidson, 1994; Lublin, 2004). While the VRA contributed to increasing Black voter registration (Davidson 1994), and eventually descriptive representation (Grose, 2011; Guinier, 1991; Lublin, 1999), gerrymandering and RPV in some localities still prevent minorities from electing their preferred candidates into office. As such, the courts are still concerned with determining whether various jurisdictions violate portions of the Voting Rights Act.

In *Thornburg v. Gingles, 478 U.S. 30, 1986*, the court established a legal framework to guide VRA challenges to legislative districts or at-large voting systems that have been accused of diluting minority voting opportunities. According to Gingles, there are three prongs that plaintiffs must establish through an analysis of voting data to make a successful claim: 1) the minority group is both geographically compact and large enough to create a single-member district; 2) the minority group tends to vote together and is politically cohesive; and 3) the non-minority (majority group) tends to vote in the opposite direction, such that it can usually block the minority groups' preferred candidate (Ross, 1993). Based on this framework and the court's prescribed statistical methods (Hood et al., 2017), social scientists were asked to employ voting analyses by relying on a combination of precinct voting data and Census block racial/ethnic data from multiple elections to assess whether a jurisdiction is in violation of the VRA.[1] At the most basic level, an analysis of ecological voting data aided the courts in answering the following important question: Do Anglos block-vote against African American candidates and prevent African Americans from gaining political representation?

Using more simple methods, the early evidence presented at trial supported what V.O. Key had already found (e.g.Goodman (1953, 1959)). Over the decades, racial demographics and social science tools have evolved considerably. King (1997) and Grofman (1992, 1995), for instance, advocated for a more precise measurement of racial voting patterns beyond homogenous precinct analysis, simple correlation techniques, and Goodman's regression. No longer facing a strictly Black-Anglo hyper-segregated environment, others, notably Rosen et al. (2001), sought ways to account for an increase in racially heterogeneous neighborhoods and the rapid emergence of Latinos and Asians.

As it stands, social scientists—and the courts—rely on two specific statistical approaches to ecological data.[2] The first, iterative ecological inference (EI), developed by King (1997), is typically preferred when there are only two racial or ethnic groups, and ideally only two candidates contesting one seat. The second and much newer and computationally intensive approach, ecological inference R x C (RxC), developed by Rosen et al. (2001), is said to be suitable when there are multiple racial or ethnic groups, or multiple candidates contesting office. While these methods make unique contributions, it is, however, unclear whether both would produce substantively different results when faced with the exact same real-world voting dataset. In one case, Grofman and Barreto (2009) used multiple ecological approaches on the same dataset and arrived at the same conclusion [for similar comparisons, also see ?]. However, others have argued that

[1]To be clear, the principal aim of the present article is not to settle the debate on the accuracy of ecological inference in the sciences writ large (e.g., see Frair et al. (2010); Freedman (1999); Greenland (2001); Martin et al. (2005); Tam Cho and Gaines (2004); Wakefield (2004), but rather to assess the degree of similarity or difference with respect to two heavily used methodologies the courts rely upon to decide whether jurisdictions are systematically discriminating against minority voters.

[2]The courts still do, however, rely on bivariate correlation, Goodman regression, and homogeneous precinct analysis. To this end, we have incorporated the Goodman regression into our R package so that analysts can assess this method alongside iterative EI and RxC.

using King's iterative EI technique with multiple racial groups or multiple candidates may produce bias estimates (Ferree, 2004). Other social scientists have gone even further, asserting in court that the iterative EI approach cannot be used to analyze multiple racial group or multiple candidate elections because "...it biases the analysis for finding racially polarized voting" (Katz, 2014).

As with any methodological advancement, there is a healthy and rigorous debate in the literature. However, very little real election data has been brought to bear. Ferree (2004) assessed King's iterative approach with simulated data and a parliamentary election in South Africa using a proportional representation system. Grofman and Barreto (2009) compared an exit poll to precinct election data in Los Angeles, but only compared Goodman's ecological regression against King's iterative EI without evaluating the RxC approach. We contribute to this literature with a comprehensive analysis of real ecological voting data from 14 elections and 78 candidates in multiethnic settings across the United States.

Using real-world ecological voting data, we aim to answer three fundamental questions not previously addressed: 1) Does the iterative EI method over-estimate racially polarized voting compared to RxC? In other words, does iterative EI bias the results towards detecting RPV? 2) Are there systematic differences in the outcomes produced by iterative EI and RxC when analyzing elections with few candidates versus elections with multiple candidates? 3) Are there systematic differences when analyzing elections with more than two racial groups?

With regards to the last two questions, if RxC is indeed a "better" method for assessing group voting behavior in a multi-candidate context as some have suggested, then one should expect to see noticeably different estimates across the two methods. Specifically, relative to RxC, the iterative EI method should become unstable and possibly generate ostensibly invalid estimates in scenarios with multiple candidates and/or multiple racial/ethnic groups. Our analysis does not find this to be the case. Instead, we find very strong patterns of consistency across iterative EI and RxC despite claims to the contrary. Across the 78 candidates we analyzed there is no evidence that either iterative EI or RxC are biased towards or against findings of polarized voting. Further, the point estimates that both methods produce are remarkably similar, typically within 2 points of one another. For social scientists and legal scholars interested in analyzing RPV when only ecological data are present, both approaches can be relied upon. Additional systematic analysis with simulated datasets leads us to the same conclusion. While our examination is fairly comprehensive and in line with other published works that compare different methods (**????**), we encourage future research to extend the bounds of our study to further examine similarities and differences between iterative EI and RxC as it pertains to RPV analysis.

In the pages that follow we first review the literature on ecological inference that is relevant to RPV analysis. Second, we describe the datasets gathered in several states spanning more than a decade. These datasets all contain elections in areas with relatively high Latino (and Anglo) voting populations and contain at least one Spanish-surnamed candidate. In addition to Latinos, many of the datasets include sizable African-American and Asian-American populations, which allows us to examine how iterative EI and RxC operate in different racial and ethnic contexts. We also examine elections with two, three, four, and up to 12 different candidates to fully assess how both models work in different electoral environments. Beyond this, we demonstrate that both the iterative EI and the RxC methods produce results in line with individual-level exit poll data. We then present Monte Carlo simulation results and introduce a congruence analysis based on a simple 2x2 comparison that can be applied to multiple groups and candidates to highlight the ways in which analysts can determine whether the two aforementioned methods result in the same substantive conclusion. Finally, we conclude with a brief discussion of our findings and some implications for the future of research in the area of ecological inference and RPV.

Ecological Inference and RPV Analysis

The challenges surrounding ecological inference are well-documented in the social science literature. Robinson (1950) pointed out that relying on aggregate data to infer the behavior of individuals can result in the ecological fallacy. Since then scholars have applied different methods to discern more accurately micro-level relationships from aggregate data. Goodman (1959; 1953) introduced ecological regression, where individual patterns can be drawn from ecological data under certain conditions. However, Goodman's statistical approach assumed that group patterns are consistent across each ecological unit, and in reality that may not be the case.

Eventually, systematic analysis revealed that early methods could produce unreliable results (see e.g., King (1997)).[3] Ecological inference is King's (1997) solution to the ecological fallacy problem inherent in aggregate data,[4] and since the late 1990s has been the benchmark method courts have relied upon to evaluate racially polarized voting patterns in voting rights lawsuits. Indeed, according to the American Constitution Society for Law and Policy, ecological inference is one of the three statistical analyses that must be performed in voting rights research on racial voting patterns.[5]

Some critics, however, have asserted that King's model was designed primarily for binary data (2x2) such as situations in which just two groups (e.g., Blacks and Anglos; Hispanics and Anglos, etc.) exist. While many geographic areas (e.g., Mississippi, Alabama) still contain essentially two groups, the growth of ethnic/racial groups such as

[3]However, in an extensive review, Owen and Grofman (1997) concluded that despite some valid theoretical concerns, the single-equation ecological regression still holds up and provides meaningful and accurate estimates of racially polarized voting. A decade later, Grofman and Barreto (2009) evaluated how ecological models compare to one another using a combination of simulated data, actual election precinct data, and an accompanying exit poll. Their analysis demonstrated that there is general consistency across the single and double equation methods, and that once voter turnout rates are accounted for similar conclusions are reached.

[4]It should be noted that ecological inference has come under criticism, especially in the fields of biological sciences, ecology, epidemiology, and public health. As Freedman (1999) has explained, when compared to available individual-level results, ecological inference estimates in epidemiology have been shown to be unreliable. In the field of ecology, Martin et al. (2005) have demonstrated that ecological techniques can lead to incorrect inferences due to the problem of zero-inflation in studies that account for the presence or absence of specific species of different animals. Greenland (2001) has outlined various potential pitfalls of ecological inference in public health research due to the non-randomization of social context across ecological units of analysis. Relatedly, Frair et al. (2010) have argued that while some ecological analysis can be informative when studying animal habitat preference, existing methods of ecological inference provide imprecise information on variation in the outcome variables, and that considerable improvements are necessary. Finally, Wakefield (2004) has too demonstrated the limits of ecological inference, especially as it pertains to questions epidemiologists are most concerned with. However, within the narrow subfield of racial voting patterns in American elections, ecological inference is still heavily relied upon, particularly by the courts.

[5]http://www.acslaw.org/sites/default/files/VRIGuidetoSection2Litigation.pdf

Latinos and Asians have challenged the historical biracial focus on race in the U.S. (Passel et al., 2011). To account for such complexities, Rosen et al. (2001) developed a hierarchical rows by columns (RxC) approach, which can be used to analyze multiple racial groups and multiple candidates together. However, due to the computationally intensive nature of their model, this approach was not initially employed in the social sciences, in general, and in voting rights cases in specific. In addition to this, King also suggested that his method can still be used with more complex data (e.g., 3x2) by "iteratively" applying the model to different subsets of the data. In trying to assess voting patterns for three racial groups (Anglos, Blacks and Hispanics), the iterative technique would estimate three separate equations. First, Anglo and Black turnout in a given electoral jurisdiction would be collapsed into a single category to estimate Hispanic vote choice for X candidate. Then, Anglos and Hispanics are grouped together to estimate X candidate support for Blacks. And finally, Hispanics and Blacks are collapsed into a single group versus Anglos to estimate X candidate support for Anglo voters.

While this iterative technique has been widely used in voting rights cases, some social scientists have expressed concern. Ferree (2004), for instance, has argued that combining Blacks and Anglos into a single "non-Hispanic" category in order to estimate Hispanic turnout may overestimate Hispanic turnout due to issues of aggregation bias and multimodality in the data. This suggests that the iterative approach could increase the likelihood of detecting racially polarized voting due to a larger-than-reality share of Hispanics in the data. While Ferree (2004) suggested some quick "fixes"—such as accounting for the relative size of each group or changing the order in which cells are estimated—to reduce aggregation bias and multimodality caused by collapsing rows or columns, she recommended estimating the cells of the rows by columns simultaneously rather than iteratively.[6] Others, such as Herron and Shotts (2003a,b), have criticized EI estimates when used for second-stage regression, given that the error is baked into the second-level regression estimation.[7] Some have gone even further in arguing King's iterative approach

[6]The simultaneous method recommended is Rosen et al.'s (2001) RxC method.

[7]In response to this issue, Adolph and King (2003) adjusted the EI procedure to reduce inconsistencies when estimating second-stage regressions.

can be "problematic and no valid statistical inferences can be drawn," and that only the hierarchical RxC approach developed by Rosen et al. (2001) can produce reliable estimates in multi-ethnic and multi-candidate settings (Katz, 2014).[8] In explaining the reasons of why the iterative EI technique is "ill-equipped" to handle complex datasets, Katz stated that "... adding additional groups and vote choices to King's (1997) EI is not straightforward," and that "... given the estimation uncertainty, it may not be possible to infer which candidate is preferred by members of the group." The argument against King's iterative EI in the case of multiple racial group, or especially multiple candidate elections, is that EI pits candidate A versus all others who are not candidate A. If the election features four candidates (A, B, C, D), critics suggest that EI cannot accurately estimate vote choice quantities because vote for candidate A is compared against the combined vote for B, C, and D. Since the iterative approach would have to run four separate equations to obtain vote estimates for each candidate, social scientists such as Katz (2014) have even claimed in court that EI biases the findings in favor of bloc-voting: "... this jerry rigged approach to dealing with more than two vote choices stacks the deck in favor of finding statistical evidence for racially polarized."

Due to these concerns, advancements in computing power, and the availability of numerous packages developed for R, the computationally intensive RxC approach is now being recommended by some in place of the iterative EI. However, no study has empirically examined how these approaches perform side-by-side with real election data containing a number of different candidate and racial group combinations. Previous work has mostly leveraged Monte Carlo simulation or only a few election datasets (**?**). Since we lack more expansive efforts to compare the two approaches, there simply is not enough information to enable researchers and legal practitioners to evaluate under which conditions the RxC method is more suitable or appropriate than the iterative EI technique. For example, if there are three racial groups in equal thirds of the electorate, does aggregation bias create

[8] Greiner and Quinn (2010) combined RxC with individual-level exit poll data, and showed that a hybrid model is perhaps even more preferable than a straight aggregation model. However, using exit poll data is not always available to researchers and practitioners. Indeed, in most county or city elections, exit poll data do not exist, which is why scholars often attempt to infer voting patterns with aggregate data.

more error in the iterative EI than a scenario in which two dominant groups comprise 90% and a small group just 10% of the electorate? Likewise, is EI's iterative approach (e.g., Black vs. non-Black, Anglo vs. non-Anglo, Hispanic vs. non-Hispanic) to candidates more stable (e.g., instability might occur if the combined vote among Blacks, Anglos, and Hispanics for just one candidate reaches over 100 percent) when analyzing three candidates and far less stable when eight candidates contest the election? Is it really the case that the iterative approach is more likely than the RxC method to produce findings in favor of racially polarized voting patterns? The analytical task of this paper is to consider these questions empirically; to systematically assess whether using the iterative EI method, as opposed to the hierarchical RxC method, can change the *substantive* conclusions one draws as it pertains to racially polarized voting patterns. Since we take advantage of real-world election datasets of varying electoral units and sizes, candidates, and racial/ethnic groups that the courts would consider, our study provides the most comprehensive attempt to answer some of the preceding questions.

Data and Methods

We turn to precinct voting data from three diverse states—California, Texas, and Florida—across 14 different elections from 2004 to 2012, in which a total of 78 candidates were on the ballot, to examine how the two different methods process the same datasets. For each of the 14 elections we analyze, we have precinct-level data on candidate vote distribution, as well as the racial demographics of the voting population in each precinct, and the total numbers of ballots cast. In two states, California and Florida, we have data on the actual voters by race and ethnicity. In Texas, we have the number of eligible voters by race and ethnicity. Thus, the key variables are percent [candidate] and percent racial/ethnic group, and our estimates control for the number of total voters per precinct, as instructed by King (1997), Ferree (2004), and Rosen and colleagues (2001).

The data we examine is diverse across almost any dimension as is illustrated by Table 1. We have data on more than 4,900 precincts in Los Angeles County or only 38 precincts in one school board district in central Florida.[9] The elections we examine also have varying number of candidates: from a head-to-head matchup with only two candidates to elections with up to twelve candidates. The data are also diverse with respect to the number of racial or ethnic groups within the electorate, starting with jurisdictions that are primarily Latino-Anglo, then areas with sizable Latino, Anglo, and Asian voting populations, and other geographies such as elections with Latino, Anglo, Asian and Black voting populations. Thus, the data we bring to bear is comprehensive and diverse across almost any metric, enabling us to follow a pattern of increasing complexity.

[TABLE 1 ABOUT HERE]

We begin the analysis with a basic dataset with just two candidates and just two racial groups, and then stick with these two racial groups and add election contests with three, four, five, six, seven, nine and twelve candidates. In each election we analyze, there is at least one co-ethnic candidate, which allows us to assess racially polarized voting patterns. After comparing Iterative EI and RxC results with two racial groups and multiple candidates, we next turn to the analysis of multiple racial groups. We first assess only two candidates, but in two different environments with Latino, Anglo and Asian, and then Latino, Anglo and Black. Then we look at both multiethnic scenarios and contests with more than two candidates. Finally, we assess a very diverse electoral environment to really put the two methods to the test. We conclude with an analysis of a Democratic primary in Los Angeles County that featured seven candidates including viable Latino, Anglo, Black and Asian candidates, and provide results for all four racial groups of voters.

Before we proceed to the election data results, it is important to briefly underscore an important issue that researchers face when dealing with aggregate-level data given that there are no bullet-proof solutions to the problem of ecological inference. Specifically, difficulties with calculating correct standard errors can arise if the aggregate data are

[9] We selected elections where we could be sure our estimates would not suffer greatly from precinct sample size concerns. However, analysts often assess racially polarized voting claims in jurisdictions with relatively few precincts. It is possible that the gaps between EI and RxC might widen with smaller precinct sample size jurisdictions, as results become increasingly unstable.

not "informative" concerning the underlying microlevel data as detailed by Tam Cho and Gaines (2004). We emphasize this particular point to not only highlight the potential pitfalls of ecological inference under certain conditions, which social scientists and legal practitioners should be aware of, but to also make the case that both iterative EI and RxC face similar constraints. That is, if a dataset is "uninformative," both approaches will suffer and produce unreliable standard errors. Conversely, if a dataset is amenable to ecological inference (i.e., meets various model assumptions), both approaches will produce relatively accurate standard errors. Therefore, under both scenarios, a side-by-side comparison of the two approaches will result in drawing identical conclusions, albeit not necessarily accurate ones depending on the dataset under consideration.

To gauge the level of information contained in a dataset, it is recommended to examine tomography plots.[10] There are two specific diagnostic uses for tomography plots. By plotting all the logically possible pairs of parameter values—that is, the known information—tomography lines can succinctly display how constrained or flexible the parameters are and thus, how difficult or easy the estimation problem will be. In a given plot, there is one tomography line bound with the [0,1] interval for each observation. Lines that do not extend across the entire unit square are further bounded than those that cross the entire unit square. If the lines are more bounded, one may be more successful when estimating the true parameter values (Tam Cho and Gaines, 2004).

In addition to showing all the available deterministic information in a problem, tomography plots also help assess whether the underlying truncated bivariate normal (TBVN) distribution imposed by ecological inference is reasonable. If most of the tomography lines seem to intersect in a region, one may conclude that the actual individual-level data are most likely, but not certainly, clustered there, marking a potential location for the mode of the joint distribution of β's. However, if no area of intersection is evident and the parameter bounds are too wide, the implication is that the TBVN distributional assumption may not be entirely met. Stated differently, if the tomography plot is considered *"uninformative,"* the data is less likely to have been generated from a TBVN

[10]Note here that as the number of parameters increase, tomography plots will become very difficult to analyze and thus, lose their diagnostic value.

distribution. This results in standard errors that may be too large to be useful or simply incorrect since they are computed based on the distributional assumption of the model (King, 1997).

When using a tomography plot, it is important to keep in mind that the information obtained from this diagnostic plot is only suggestive. A tomography plot does not allow a researcher to make definitive claims about the particular distributional assumptions of the data. As Tam Cho and Gaines (2004) have stated, "...deciding whether a tomography plot is informative is something of an art, no one has devised a concrete measure for *'informativeness'* or any formal test for accepting or rejecting the TBVN distributional assumption (or any other distributional assumption) on the basis of the plot" (pg. 155). What this means is that tomography plots only provide an indication of the risk associated with forcing a distributional assumption on the data. If the parameter bounds are too wide and there is no general area of intersection, incorrect standard errors may be obtained (King, 1997).

Despite the challenges that one faces when analyzing tomography plots, especially as the number of parameters increase, such inspection is worthwhile in helping researchers evaluate the extent to which the ultimate conditional distributions are fairly close approximations to the truth. If tomography plots lead one to reject the TBVN distributional assumption, the ecological inference method may still be appropriate if one conditions on suitable covariates (Tam Cho and Gaines, 2004).[11]

Our assessment of tomography plots suggests that some datasets are certainly more "informative" than others. For example, Figure 1 demonstrates examples where the tomography lines tend to intersect in one general area, and the parameter bounds are fairly narrow in that they do not extend across the entire unit square of the plot. Based on these plots, one can make a reasonable case that the data has been generated from a TBVN distribution. In contrast, Figure 2 displays tomorgraphy plots that are considered less informative because the lines intersect in multiple areas or/and the parameter bounds are fairly wide. In cases in which the tomography plot indicates other distributional

[11]Therefore, tomography plots can also be viewed as a diagnostic tool for determining the necessity of adding appropriate covariates to the model.

assumptions, the standard errors that one obtains is likely, but not certainly, inaccurate.

The discussion surrounding the distributional assumptions imposed by ecological inference leads to our key point: if the datasets are not consistent with the specified TBVN distribution, neither interative EI or EI RxC (or even OLS) will produce accurate standard errors unless one introduces relevant covariates into the model (Tam Cho and Gaines, 2004). Thus, one cannot, on the basis of such diagnostics, make the claim that the RxC approach, which faces similar constraints as the iterative approach, somehow produces more or less accurate estimates. As the forthcoming results will demonstrate, a comparison of the two approaches yields similar substantive conclusions about the presence or absence of racially polarized voting regardless of the varying degrees of estimation difficulty.

[FIGURE 1 ABOUT HERE]

[FIGURE 2 ABOUT HERE]

Election Data Results

Using the R packages ei (King and Roberts, 2012) and eiPack (Lau et al., 2006), we estimated vote choice for candidates across racial groups using precinct-level election data.[12] For EI, we took the iterative approach that has been questioned by some. In this approach, we iteratively estimated how each racial group voted for each candidate. So in an election with three different racial groups and seven different candidates we estimated a total of 21 EI models. In contrast, the RxC approach allows analysts to estimate all the models simultaneously. Recall, our overarching question is: Does the iterative approach over-estimate racially polarized voting (RPV) compared to the RxC approach?

Despite various claims regarding the potential limitations of the iterative approach, we find no statistically different vote estimates across the 14 elections and 78 candidates we examined in the EI versus RxC approach (all the results race by race and candidate

[12]We have written our own package combining these packages, which will be available on CRAN and Github upon article acceptance. All of our datasets and replication code will also be made publicly available

by candidate can be seen in the Appendix tables 12 to 24. Simply stated, our analysis reveals that both methods lead to substantively similar conclusions about vote choice and racially polarized voting.

Where differences do emerge, there is no consistent pattern in whether EI or RxC produce higher or lower levels of racially polarized voting, contrary to some assertions. In some instances EI might yield 1 point higher minority vote cohesion, but in other instances RxC estimates 2 points higher minority vote cohesion, and in every instance the minority vote estimates are statistically indistinguishable from one another. Overall, we estimated 193 racial group-candidate vote outcomes and found that in 73 percent of the cases the difference between EI and RxC is within only 2 points. More specifically, in 105 instances the difference in the vote choice estimate is less than 1 point, and in 35 instances the difference is between 1 to 2 points. This suggests remarkable consistency across the two approaches. For the remaining 27 percent of the cases, only 11 of them—or 6 percent—produced estimates that were over 5 points different from one another, as summarized in Table 2.

[TABLE 2 ABOUT HERE]

We also found no evidence that EI is more likely to produce results in favor of racially polarized voting. For example, in the first election we considered, EI reports slightly higher minority cohesion—84.04 (EI) to 82.94 (RxC)—for the Latino-preferred candidate. However, in the second election we examined RxC reports slightly higher minority cohesion—94.39 (EI) to 96.56 (RxC)—for the Latino preferred candidate. In 20 instances in which minority voters had a minority preferred candidate, EI produced higher minority cohesion 8 times and RxC produced higher minority cohesion 12 times (see table 3). In some instances this difference in "higher cohesion" amounts to less than a half-point difference such as the Latino candidate, Torrico, winning an estimated 18.24 percent of the Latino vote in RxC versus an estimated 17.85 percent under EI. Thus, even where differences emereged they were often negligible and would round to the same whole number. Likewise, we found no evidence that Anglo bloc voting against minority-preferred candidates is stronger under EI as compared to RxC, with each method sometimes producing slightly higher Anglo bloc voting exactly half of the time (see Table 3).

[TABLE 3 ABOUT HERE]

Recall that our second research questions was: are there systematic outcome differences between EI and RxC when analyzing elections with few candidates versus elections with multiple candidates? We might expect greater differences to emerge when there are more candidates than fewer candidates—the claim is that RxC is designed for this scenario whereas EI is more equipped in dealing with 2x2 datasets. Another way of stating this is: Do EI and RxC essentially produce the same results when there are two, or maybe three candidates, but start to diverge when six, seven or more than ten candidates are on the ballot?

In the first section of our analysis we compared EI and RxC with only two racial groups—Latinos and Anglos—across eight elections in which the number of candidates on the ballot varied from two to twelve. The elections consisted of contests in Los Angeles, CA; Orange County, CA; Corona, CA; Orange County, FL; Oceanside, CA; Vista, CA; and San Mateo, CA. This diversity allowed us to assess whether the number of candidates impacted the stability of EI and RxC estimates. Table 4 shows the co-ethnic minority preferred candidate for each one of the eight elections. Figure 3 visualizes the differences between method estimates by race for each election. As is illustrated, there is no detectable pattern that would lead one to conclude that the iterative EI is more likely to produce results in favor of racially polarized voting. Furthermore, even when the datasets were more or less amenable to ecological inference based on an assessment of tomography plots, the conclusions regarding racially polarized voting did not change. For instance, in the Cardona and Vista election results, the datasets were considered more "informative" in that parameter bounds were relatively narrow and a general area of intersection existed. In contrast, the Los Angeles and San Mateo elections were cases in which the datasets were considered less informative. Nevertheless, both approaches produced similar outcomes. That is, no patterns were detected with more or less informative datasets given that both methods face similar estimation constraints if certain conditions, such as the TBVN distributional assumption, are not met.

[TABLE 4 ABOUT HERE]

[FIGURE 3 ABOUT HERE]

So far we have only examined races with two racial groups (Latino and Anglo). In the next section we compare EI and RxC in six elections with more than two racial groups; two elections with Latinos, Asians, and Anglos; three with Latinos, Blacks, and Anglos; and one election with the four racial groups. This allows us to assess our third major question: Are there systematic outcome differences between EI and RxC when analyzing elections with more than two racial groups?

In addition to examining elections with different racial group combinations, our data enabled us to consider elections with as low as two and as high as twelve candidates so that we can continue to assess whether systematic differences emerge between EI and RxC in much more complex environments. Tables 5, 6, 7 report the co-ethnic minority preferred candidate for each one of the elections examined. Similarly, Figures 4 and 5 visualize the differences. Finally, Figure 6 presents a compiled visualization of all the races with more than two ethnic groups. The results display remarkable similarity between EI and RxC estimates even as the number of ethnic groups and candidates increase. Once again, we did not detect any patterns that would lead one to conclude that EI is more or less likely than RxC to produce results in favor of racially polarized voting.

[TABLE 5 ABOUT HERE]

[FIGURE 4 ABOUT HERE]

[FIGURE 5 ABOUT HERE]

[TABLE 6 ABOUT HERE]

[FIGURE 6 ABOUT HERE]

[TABLE 7 ABOUT HERE]

Comparison with Exit Poll Data

In many, if not most, situations where analysts are called to evaluate the presence or absence of racially polarized voting, ecological inference is the chosen method in part because individual-level polling data are unavailable. For instance, pollsters do not collect data for elections in small cities, such as Blythe, CA. In major cities, though, exit poll data are occasionally available.

While our main question is whether EI and RxC produce substantively similar RPV outcomes, there is a possibility that EI may be inaccurate relative to the "truth" more often than the RxC approach. To consider this possibility, we compare EI and RxC estimates in a voting scenario with known outcomes that provide vote choice by race/ethnicity (i.e., an exit poll or pre-election poll). To be sure, exit polls can produce biased estimates of subgroups because of the reliance on "bellwether" counties or precincts comprising heterogenous populations including racial/ethnic groups (Barreto et al., 2006; Mitofsky, 1998; Traugott and Price, 1992). Specifically, Barreto et al. (2006) argue that heterogeneous precinct-based exit polls often overestimate Republican support among Latino voters because pollsters selecting bellwether precincts are more likely to encounter acculturated Latinos who are disproportionately Republican. That said, an exit poll is still another point of comparison employed to get closer to the actual individual-level voting behavior.

Many studies have pointed out that ecological fallacy and other estimation issues can produce ecological inference results that are unreliable. While we acknowledge the limitations of ecological inference, we find that the results from EI and RxC are similar to the individual-level exit poll data as it pertains to evaluating racially polarized voting patterns in Voting Rights Act cases. Table 8 presents EI and RxC estimates for the 2005 Los Angeles mayoral runoff election between Antonio Villaraigosa (Latino) and James Hahn (Anglo). These estimates are compared to results from the Los Angeles Times exit poll. Our findings demonstrate that not only do EI and RxC produce remarkably similar estimates, but that the results closely match the individual-level estimates from the Los Angeles Times poll. More specifically, the EI method estimates Villaraigosa receiving

82 percent of the Latino vote and only 45 percent of the Anglo vote; the RxC method estimates Villaraigosa receiving 81 percent of the Latino vote and just 48 percent of the Anglo vote. If the task is to evaluate a pattern of racially polarized voting, both methods closely match the conclusion one would draw from the exit poll, which reports that an estimated 84 percent of Latinos voted for Villaraigosa while only 50 percent of Anglo voters made the same choice. While the EI method shows slightly more RPV compared to the RxC method in this particular case, the difference is very substantively negligible in voting rights lawsuits. Moreover, the EI and RxC estimates are all with the margin of error of the individual-level data reported by the LA Times exit poll. In sum, this comparison provides additional evidence that both methods may be useful in evaluating RPV in Voting Rights Act cases.

[TABLE 8 ABOUT HERE]

Monte Carlo Simulation Results

While the analyses with real-world election data demonstrated congruence between the two methods, Monte Carlo simulations provide another way of evaluating our most basic question: do analysts reach substantively different conclusions when comparing iterative EI and RxC estimates?[13] To answer this question, we drew simulations from a beta distribution with parameters β=2, α=2 to construct the following datasets: 2 candidates, 2 groups; 2 candidates, 3 groups; 3 candidates, 2 groups, 3 candidates, 3 groups; and 4 candidates and 4 groups. Each dataset contains anywhere from 100 to 1000 precincts, and each precinct ranges in size from 10 to 1000 total voters. The data also contain a set of columns for each group's simulated percent share of the precinct and percent vote for the hypothetical candidates. For each of the dataset types (2x2, 2x3, 3x2, 3x3, 4x4), we then randomly generated 100 datasets, estimated group votes using both iterative EI and RxC methods, and stored the average difference between the

[13]However, we note that simulations are not necessarily a "better approach" since randomly generated data could contain many scenarios in which there are no clear minority-preferred candidates—that is, cases that are of little interest to potential plaintiffs.

two methods across all groups and candidates. Figure 7 visually depicts the simulation results.

The findings largely validate the results obtained with real-word election data. Across 500 randomly generated datasets, we find tremendous consistency between the two methods, with overall mean differences by each election type ranging between 1 to 4 percentage points. In voting rights cases, these observed differences would almost never alter one's substantive conclusions about racially polarized voting patterns. Even in the rare cases where we found larger discrepancies (e.g., only 8% of the 2x3 data types), both methods concurred on the hypothetical groups' preferred candidate. A detailed look at the results, for instance, revealed that iterative EI estimated that 80% of group 1 favored candidate 2, while RxC estimated that 90% of group 1 favored the same candidate. Thus, for all practical purposes, experts would reach similar conclusions about RPV as the two methods concurred on the direction and degree of polarization.

Finally, and perhaps most importantly, our assessment of the simulation results did not reveal any systematic patterns where iterative EI produced higher or lower estimates than the RxC method. In some cases RxC produced higher group estimates while in other cases iterative EI did.

[FIGURE 7 ABOUT HERE]

Model Congruence Score: Do the Two Methods Lead to Similar Substantive Conclusions?

The previous sections demonstrated that iterative EI and RxC tend to produce similar vote choice estimates under various conditions. However, our discussion of the "substantive" evaluation of the results did not provide a systematic way of interpreting the findings. A systematic evaluation of congruence between the two methods is important because plaintiffs must show judges that racially polarized voting (RPV) exists, and that RPV is not just a function of choosing one statistical method over another, but something that generally holds regardless of the approach. Social scientists are also similarly interested in understanding the extent to which results are substantively consistent across different estimators.

To this end, we introduce a new approach to aid analysts in determining whether the two methods produce similar judgements, which we call the Model Congruence Score (MCS). The MCS can be applied in either 2x2 settings or with some adjustments extended to situations with multiple candidates and multiple groups, although analysts may want to set some decision rules in terms of whether to combine all candidates of the same race together (e.g., one election might have multiple Anglo candidates: Smith, Toms, and Johnson) into one racial group candidate for the purpose of assessing RPV patterns.

What exactly can the MCS reveal with respect to voting right analysis? First, do both iterative EI and RxC conclude that minority voters prefer the minority candidate and that Anglo voters prefer the minority candidate? If minority voters prefer the Anglo candidate and so do Anglo voters, then RPV does not exist. Likewise, if both minority and Anglo voters both prefer the minority candidate, RPV does not exist. Both cases would not meet the Gingles threshold outlined by the court. To answer this initial question, the MCS rates whether simple polarized voting exists based on the estimates obtained from iterative EI and RxC.

Second, what is the relative degree of RPV in each of the models? For example, do both models suggest a 30-point gap in racial voting preference, or does one model suggest only a 5-point difference and the second model suggests a 40-point difference? The difference in voting preferences, and not just the direction of preferences, is a very important component of the congruence score and informative to the courts. In order to answer this second question, MCS first estimates the percentage point gap between minority and Anglo voters for the minority preferred candidate, and then for the Anglo-preferred candidate. Next, MCS evaluates what percentage of minority voters would need to switch from voting for the minority candidate to supporting the Anglo candidate such that there is an even 50-50 distribution, and no clear preferred candidate. Likewise, MCS calculates the percentage of Anglo voters that would need to switch from voting for the Anglo candidate to supporting the minority candidate to create a 50-50 distribution.

While the math is different, the logic behind this measure is similar to the dissimilarity index commonly used in demography (Massey and Denton, 1988).

Third, if voting patterns hold, are minority voters blocked by Anglo voters from electing a minority candidate? And by how much are they blocked? Again this step adds both a simple 'yes/no' distinction of being blocked, but also calculates and compares the degree by which a minority-preferred candidate is blocked. Overall, the MCS attempts to provide a simple measure, ranging from 0 to 1, to assess how much congruence exists between and within the vote choice estimates across iterative EI and RxC.

We first calculated MCS for both iterative EI and RxC in a simple 2x2 configuration to show in more detail how the process works. We report congruence scores for each metric, which is scaled from 0-1, where 0 reveals the two methods are in complete disagreement, and 1 indicates the two methods are in complete agreement. For ease of interpretation, we explain the precise metrics for the aforementioned three tests and their congruence with actual data from iterative EI and RxC estimates of the Latino and non-Latino vote from the 2010 Los Angeles County Insurance Commissioner race where the Latino candidate, De la Torre, ran against Jones (Anglo). While the non-Latino group includes non-Latino minorities, for simplicity, we bin Anglos with non-Latino minorities in order to craft a simple 2x2 scenario (see Table 11 in the appendix for full vote choice estimates).

To assess whether Latino voters prefer the Latino candidate, we examine the difference between Latino support for De la Torre and Anglo/other support for De la Torre. According to Table 11, iterative EI places Latino support for De la Torre at 84 percent, whereas for Anglo/non-Latinos the estimate is at just 22 percent. The candidate support by racial group is thus just over 62 percent, which is shown in column 2, labeled EI, the first row of Table 9. The same calculation is made for the RxC method, placing Latino support for De la Torre at about 83 percent and the non-Latino support at about 23 percent—a difference of 60 percentage points. How similar are these findings? To calculate the congruence score on this measure we take the absolute difference between the iterative EI and RxC estimate for Latino - non-Latino support for De la Torre then divide this by the absolute mean difference of the two methods. Finally, to transform this into a 0-1 scale, where 1 equals complete congruence and 0 equals no congruence, we subtract the resulting value from 1 so that values closer to 1 imply higher congruence:

$$
\begin{aligned}
&x = \text{EI Latino vote for De la Torre} - \text{EI Non-Latino vote for De la Torre}\\
&y = \text{RxC Latino vote for De la Torre} - \text{RxC Non-Latino vote for De la Torre}\\
&\quad = 1 - \frac{abs(x-y)}{abs(mean(x,y))}
\end{aligned} \tag{1}
$$

We can plug the data from Table 11 into the equation above to produce the congruence score, which is identical to the congruence score appearing on row one of Table 9:

$$
\begin{aligned}
&= 1 - \frac{abs((84.11 - 22.02) - (82.94 - 22.99))}{abs(mean((84.11 - 22.02), (82.94 - 22.99)))}\\
&= 0.965
\end{aligned} \tag{2}
$$

Row two in Table 9 assesses whether De la Torre is preferred by Latino voters. The congruence receives 1 if both the iterative EI and RxC method reveal that Latinos preferred De la Torre to Jones (or 1 if both methods revealed a preference for Jones). In the present case, both methods show that Latinos prefer De la Torre, so the congruence on this metric receives a 1. The preference rate is calculated as the difference between Latino support for the Latino candidate, De la Torre, and the Anglo candidate, Jones. For iterative EI, this would be 84.11 - 15.92. The resulting figure is then divided by 2, to show how much above the 50 percent mark De la Torre is preferred over Jones. In other words, what is the percentage of Latino voters who would have to switch to Jones so that Latinos did not prefer either candidate? For iterative EI this number is 34. Using the same calculation for RxC, we arrive at nearly 33. Thus, our numbers in this case are very similar, and so a congruence score of 0.966 is reported. The equations for this congruence are listed below:

$$
\begin{aligned}
x &= (\text{EI Latino vote for De la Torre} - \text{EI Latino vote for Jones})/2 \\
y &= (\text{RxC Latino vote for De la Torre} - \text{RxC Non-Latino vote for Jones})/2 \\
&= 1 - \frac{abs(x-y)}{abs(mean(x,y))}
\end{aligned} \tag{3}
$$

The actual numbers are presented here:

$$
\begin{aligned}
x &= (84.11 - 15.92)/2 \\
y &= (82.94 - 17.06)/2 \\
&= 1 - \frac{abs(x-y)}{abs(mean(x,y))} \\
&= 0.966
\end{aligned} \tag{4}
$$

Finally, we turn to vote blocking. Given the way districts are often drawn, this is a crucial question posed to judges who assess whether Anglos are blocking Latinos from electing their preferred candidates (usually Latino). In our working example, for non-Latinos we subtract their support for Jones from non-Latino support for De la Torre. This is then divided by two (as in the above set of equations). This essentially measures how much Anglos (or non-Latinos) support the Anglo candidate, and how many votes they would need to dole out to the Latino candidate to not block the Latino candidate from getting elected. For iterative EI, this is (22 - 78)/2, and for RxC this is (23 - 77)/2. Once again, the congruence score is calculated in a similar way as above, which produces a score of 0.965. Row four of Table 9 also reports whether Anglos are, in general, block voting against Latinos— if both the iterative EI and RxC agree, then the congruence is given a 1.

$$
\begin{aligned}
x &= (\text{EI Non-Latino vote for De la Torre} - \text{EI Non-Latino vote for Jones})/2 \\
y &= (\text{RxC Non-Latino vote for De la Torre} - \text{RxC Non-Latino vote for Jones})/2 \\
&= 1 - \frac{abs(x-y)}{abs(mean(x,y))}
\end{aligned}
\tag{5}
$$

$$
\begin{aligned}
x &= (22.02 - 77.99)/2 \\
y &= (22.98 - 77.01)/2 \\
&= 1 - \frac{abs(x-y)}{abs(mean(x,y))} \\
&= 0.965
\end{aligned}
\tag{6}
$$

For the total Latino candidate congruence score, we take the mean of the existing congruence scores, resulting in a final score of 0.979. The process is reversed for calculating the requisite scores for the Anglo candidate. In the 2x2 scenario, the numbers are essentially the same as those calculated for the minority candidate; however the coefficient sign is switched, and the block rate and preference rates are swapped. The final step taken to obtain an "overall" or "total model congruence score" is to then calculate the average of the minority and Anglo candidate congruence scores obtained in the previous steps (see result below).

[TABLE 9 ABOUT HERE]

Beyond the 2x2 example, we also provide detailed model congruence scores in the appendix for a 2x4, 2x5, 3x2, and 4x7 election analysis comparing iterative EI and RxC. For ease of interpretation, Table 10 summarizes the overall congruence scores for all elections analyzed. Overall, the findings demonstrate high levels of congruence across a variety of different elections with multiple candidates and multiple racial/ethnic voter groups.

[TABLE 10 ABOUT HERE]

Conclusion

This paper engages an important methodological topic with real-world implications. Specifically, we examined three questions to assist social scientists, legal practitioners, and the courts working with Voting Rights Act cases in which only aggregate-level (e.g., precinct) data exist: 1) Does ecological inference's (EI) iterative technique over-estimate racially polarized voting (RPV) compared to RxC? In other words, does EI bias towards detecting RPV? 2) Are there systematic outcome differences between EI and RxC when analyzing elections with few candidates versus elections with multiple candidates? 3) Are there systematic outcome differences between iterative EI and RxC when analyzing elections with more than two racial groups? These questions were assessed with real-world data from 14 elections with 78 candidates and 500 simulated datasets of varying number of candidates and groups.

To examine whether voting districts experienced RPV, we estimated vote shares for different candidates from voters of different racial/ethnic groups using two of the most commonly used ecological inference methods. We evaluated King's iterative ecological inference (EI) approach against the more recent rows by columns (RxC) approach. Using elections with multiple candidates and multiple groups (i.e., Latinos, Anglos, Blacks, Asians), we did not find significant differences between the two methods in terms of estimating candidate support. To the extent that differences did emerge, they were not systematic in any way. Furthermore, in one analysis where exit poll data were available, we compared the iterative EI and RxC results against known exit poll figures and found that the three methods produced statistically and substantively indistinguishable candidate estimates for different racial/ethnic voting blocs. A series of Monte Carlo simulations provided additional support for the assertion that iterative EI and RxC produce substantively similar estimates in different candidate-group combinations.

Finally, we presented a new congruence test that analysts can implement to interpret RPV patterns when using both iterative EI and RxC methods. We outlined how analysts can calculate model congruence scores (MCS) ranging from 0-1, where 0 indicates iterative EI and RxC produce completely opposite results, and 1 indicates that the methods are in complete agreement. We then applied this test to a host of elections, finding that overall congruence between the two techniques is very high. In other words, a MCS analysis provides a quantitative figure to assess EI/RxC congruence; in the present scenario these figures suggest no meaningful differences between the two methods. To our knowledge, this is the first usage of MCS, and one of the key contributions of this paper.

Our findings have important implications for academics and practitioners who evaluate litigation in the voting rights arena. While there has been a robust debate on precisely what method to use, we suggest that claims about the superiority of one method over the other should not be made without clear and convincing evidence. While we find no concerning discrepancies between the two methods in the elections we analyzed, we do not claim that our analysis rests all debate. Rather, we invite social scientists to further examine the weaknesses and strengths of different approaches as they pertain to identifying the presence or lack of racially polarized voting patterns.[14]

[14]To this end, we posted our R package to CRAN so that other researchers can employ similar analyses. Data and replication code will also be posted online.

References

Christopher Adolph and Gary King. Analyzing second-stage ecological regressions: Comment on herron and shotts. *Political Analysis*, 11(1):65–76, 2003.

Matt A Barreto. Si se puede! latino candidates and the mobilization of latino voters. *American Political Science Review*, 101(3):425–441, 2007.

Matt A Barreto. *Ethnic cues: The role of shared ethnicity in Latino political participation.* University of Michigan Press, 2010.

Matt A Barreto, Mario Villarreal, and Nathan D Woods. Metropolitan latino political behavior: Voter turnout and candidate preference in los angeles. *Journal of Urban Affairs*, 27(1):71–91, 2005.

Matt A Barreto, Fernando Guerra, Mara Marks, Stephen A Nuño, and Nathan D Woods. Controversies in exit polling: Implementing a racially stratified homogenous precinct approach. *PS: Political Science & Politics*, 39(3):477–483, 2006.

Adam B Cox and Thomas J Miles. Judging the voting rights act. *Columbia Law Review*, pages 1–54, 2008.

Chandler Davidson. *Quiet revolution in the South: The impact of the Voting Rights Act, 1965-1990.* Princeton University Press, 1994.

Michael C Dawson. *Black visions: The roots of contemporary African-American political ideologies.* University of Chicago Press, 2003.

Karen E Ferree. Iterative approaches to rxc ecological inference problems: where they can go wrong and one quick fix. *Political Analysis*, 12(2):143–159, 2004.

Jacqueline L Frair, John Fieberg, Mark Hebblewhite, Francesca Cagnacci, Nicholas J DeCesare, and Luca Pedrotti. Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1550):2187–2200, 2010.

David A Freedman. Ecological inference and the ecological fallacy. *International Encyclopedia of the social & Behavioral sciences*, 6:4027–4030, 1999.

Leo A Goodman. Ecological regressions and behavior of individuals. *American sociological review*, 1953.

Leo A Goodman. Some alternatives to ecological correlation. *American Journal of Sociology*, pages 610–625, 1959.

Sander Greenland. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *International journal of epidemiology*, 30(6):1343–1350, 2001.

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

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

Bernard Grofman. The use of ecological regression to estimate racial bloc voting. *UsFL rev.*, 27:593, 1992.

Bernard Grofman. New methods for valid ecological inference. *Spatial and Contextual Models in Political Research*, pages 127–149, 1995.

Bernard Grofman and Matt A Barreto. A reply to zax's (2002) critique of grofman and migalski (1988) double-equation approaches to ecological inference when the independent variable is misspecified. *Sociological Methods & Research*, 37(4):599–617, 2009.

Bernard Grofman and Lisa Handley. Minority population proportion and black and hispanic congressional success in the 1970s and 1980s. *American Politics Quarterly*, 17 (4):436–445, 1989.

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

Christian R Grose. *Congress in black and white: Race and representation in Washington and at home*. Cambridge University Press, 2011.

Lani Guinier. The triumph of tokenism: The voting rights act and the theory of black electoral success. *Michigan Law Review*, 89(5):1077–1154, 1991. ISSN 00262234. URL http://www.jstor.org/stable/1289550.

Michael C Herron and Kenneth W Shotts. Cross-contamination in ei-r: Reply. *Political Analysis*, 11(1):77–85, 2003a.

Michael C Herron and Kenneth W Shotts. Using ecological inference point estimates as dependent variables in second-stage linear regressions. *Political Analysis*, 11(1):44–64, 2003b.

MV Hood, Peter A Morrison, and Thomas M Bryan. From legal theory to practical application: A how-to for performing vote dilution analyses. *Social Science Quarterly*, 2017.

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

Jonathan N. Katz. Expert report on voting in the city of Whittier. Superior Court of the State of California, March 2014.

Vladimer Key. Southern politics in state and nation. 1949.

Gary King. A solution to the ecological inference problem, 1997.

Gary King and Margaret Roberts. Ei: a(n) r program for ecological inference. *Harvard University. Retrieved from http://gking.harvard.edu/files/ei.pdf*, 2012.

Olivia Lau, Ryan T Moore, and Michael Kellermann. eipack: Rxc ecological inference and higher-dimension data management. *New Functions for Multivariate Analysis*, 18 (1):43, 2006.

David Lublin. *The paradox of representation: Racial gerrymandering and minority interests in Congress.* Princeton University Press, 1999.

David Lublin. *The republican south: democratization and partisan change.* Princeton University Press, 2004.

Tara G Martin, Brendan A Wintle, Jonathan R Rhodes, Petra M Kuhnert, Scott A Field, Samantha J Low-Choy, Andrew J Tyre, and Hugh P Possingham. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology letters*, 8(11):1235–1246, 2005.

Douglas S Massey and Nancy A Denton. The dimensions of residential segregation. *Social forces*, 67(2):281–315, 1988.

Peyton McCrary. Racially polarized voting in the south: Quantitative evidence from the courtroom. *Social Science History*, 14(04):507–531, 1990.

Warren J Mitofsky. Was 1996 a worse year for polls than 1948? *The Public Opinion Quarterly*, 62(2):230–249, 1998.

Guillermo Owen and Bernard Grofman. Estimating the likelihood of fallacious ecological inference: linear ecological regression in the presence of context effects. *Political Geography*, 16(8):675–690, 1997.

Jeffrey S Passel, D v Cohn, and Mark Hugo Lopez. Hispanics account for more than half of nation?s growth in past decade. *Washington, DC: Pew Hispanic Center*, 2011.

Spencer Piston. How explicit racial prejudice hurt obama in the 2008 election. *Political Behavior*, 32(4):431–451, 2010.

William S Robinson. Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3), 1950.

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

Mary Massaron Ross. The voting rights act. *Urb. Law.*, 25:925, 1993.

Wendy K Tam Cho and Brian J Gaines. The limits of ecological inference: The case of split-ticket voting. *American Journal of Political Science*, 48(1):152–171, 2004.

Katherine Tate. *From protest to politics: The new black voters in American elections.* Harvard University Press, 1994.

Michael W Traugott and Vincent Price. The polls?a review: Exit polls in the 1989 virginia gubernatorial race: Where did they go wrong? *Public Opinion Quarterly*, 56 (2):245–253, 1992.

Jon Wakefield. Ecological inference for 2x2 tables (with discussion). *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 167(3):385–445, 2004.

Tables[15]

Table 1: Summary Table of Elections Analyzed

Geography	Year	Ethnic Grps	# Cand.	Contest	Precincts
Los Angeles Co., CA	2010	2 (L, W)	2	Insurance Commissioner Dem Primary	4,980
Orange Co., FL	2006	2 (L, W)	3	School Board	44
Corona, CA	2006	2 (L, W)	4	City Council	47
Orange Co., FL	2012	2 (L, W)	5	County Commission	38
Corona, CA	2004	2 (L, W)	6	City Council	48
Oceanside, CA	2012	2 (L, W)	7	City Council	78
Vista, CA	2012	2 (L, W)	9	City Council	36
San Mateo, CA	2010	2 (L, W)	12	Superintendent of Public Education	433
Orange Co., CA	2010	3 (L, W, A)	2	Insurance Commissioner Dem Primary	1,941
Fullerton, CA	2012	3 (L, W, A)	12	City Council	84
Harris Co., TX	2010	3 (L, W, B)	2	Land Commissioner	885
Harris Co., TX	2010	3 (L, W, B)	3	Lieutenant Governor Dem Primary	885
Orange Co., FL	2008	3 (L, W, B)	4	Soil & Water Board of Directors	252
Los Angeles Co., CA	2010	A (L, W, B, A)	7	Attorney General Dem Primary	4,974

Note: L= Latino, W=White, B=Black, A=Asian

[15] We use the term white to mean Anglo in all tables and figures.

Table 2: Distribution of difference between EI and RxC vote choice estimates

EI vs. RxC outcome	n	%
Less than 1 point difference	105	54%
1 to 2 points difference	35	18%
2 to 3 points difference	19	10%
3 to 4 points difference	8	4%
4 to 5 points difference	15	8%
Over 5 points difference	11	6%

Out of 193 vote choice scenarios

Table 3: Comparison of which method produces stronger racially polarized voting estimates in conditions with minority-preferred candidate

	Minority cohesion	White bloc voting
EI stronger polarization	8	10
RxC stronger polarizaton	12	10

Out of 20 instances where minority voters had a minority preferred candidate

Table 4: Elections with 2 Ethnic Groups (Latino & White)

Geography	# of Candidates	EI vs RxC estimate difference Latinos	Whites
Los Angeles Co., CA	2	-1.10	1.04
Orange Co., FL	3	2.17	-0.75
Corona, CA	4	-0.96	0.29
Orange Co., FL	5	2.78	-0.73
Corona, CA	6	0.76	-0.11
Oceanside, CA	7	-4.52	0.93
Vista, CA	9	1.05	-0.31
San Mateo, CA	12	-1.32	0.16

Table 5: Elections with 3 Ethnic Groups (Latino Blacks, & White)

	# of	EI vs RxC estimate difference		
Geography	Candidates	Latinos	Whites	Blacks
Harris CO, TX	2	-4.62	-8.59	4.63
Orange Co., FL	4	0.14	-1.20	-3.76
Harris CO, TX	3	0.01	1.73	-4.65

Table 6: Elections with 3 Ethnic Groups (Latino, Asian & White)

	# of	EI vs RxC estimate difference		
Geography	Candidates	Latinos	Whites	Asians
Orange Co., CA	2	2.95	-0.90	-6.78
Fullerton, CA	12	1.72	-0.80	2.79

Table 7: Elections with 4 Ethnic Groups (Latino, Black, Asian, & White)

	# of	EI vs RxC estimate difference			
Geography	Candidates	Latinos	Whites	Asians	Blacks
Los Angeles Co., CA	7	1.23	1.19	2.54	-2.85

Table 8: Percent voting for Antonio Villaraigosa (AV) and James Hahn (JH) by ethnic group. Comparison between EI, RxC, and exit poll methods, Los Angeles mayoral election runoff, May 2005. Exit poll taken from Los Angeles Times.

	EI: AV	EI: JH	RxC: AV	RxC: JH	Exit: AV	Exit: JH	MOE
White	45	54	48	52	50	50	+/- 2.5
Black	58	40	50	50	48	52	+/-4.2
Latino	82	17	81	19	84	16	+/-3.6
Asian	48	51	47	53	44	56	+/-6.1

Table 9: 2x2 Congruence table for Los Angeles County Insurance Commissioner Election 2010

	EI	RxC	Congruence
MV1-WV for MC1	62.091	59.955	0.965
MC1 preferred by MV1	Yes	Yes	1
MC1 preference rate	34.094	32.946	0.966
MC1 blocked by WV	Yes	Yes	1
MC1 block rate	-27.986	-27.012	0.965
MC Model Congruence			0.9792
MV1-WV for WC1	-62.091	-59.955	0.965
WC1 preferred by WV1	Yes	Yes	1
WC1 preference rate	27.986	27.012	0.965
WC1 blocked by MV1	Yes	Yes	1
WC1 block rate	-34.094	-32.946	0.966
WC Model Congruence			0.9792
Total Model Congruence Score			0.9792

Note: see Table 11 for actual polarized voting results for EI and RxC

Table 10: Summary of overall model congruence scores across all elections analyzed

RxC	Georgraphy	Precinct (n)	Congruence
2x2	Los Angeles, CA	4980	0.9792
2x3	Orange County, FL	44	0.9818
2x4	Corona, CA	47	0.9033
2x5	Orange County, FL	38	0.8829
2x6	Corona, CA	48	0.8546
2x7	Oceanside, CA	78	0.7857
2x9	Vista, CA	36	0.9377
2x12	San Mateo, CA	433	0.9561
3x2	Orange County, CA	1941	0.8169
3x12	Fullerton, CA	84	0.8344
3x2	Harris County, TX	885	0.9081
3x3	Harris County, TX	885	0.7952
3x4	Orange County, FL	252	0.8695
4x7	Los Angeles, CA	4974	0.8717

Figures

Figure 1: More "Informative" Tomography Plots

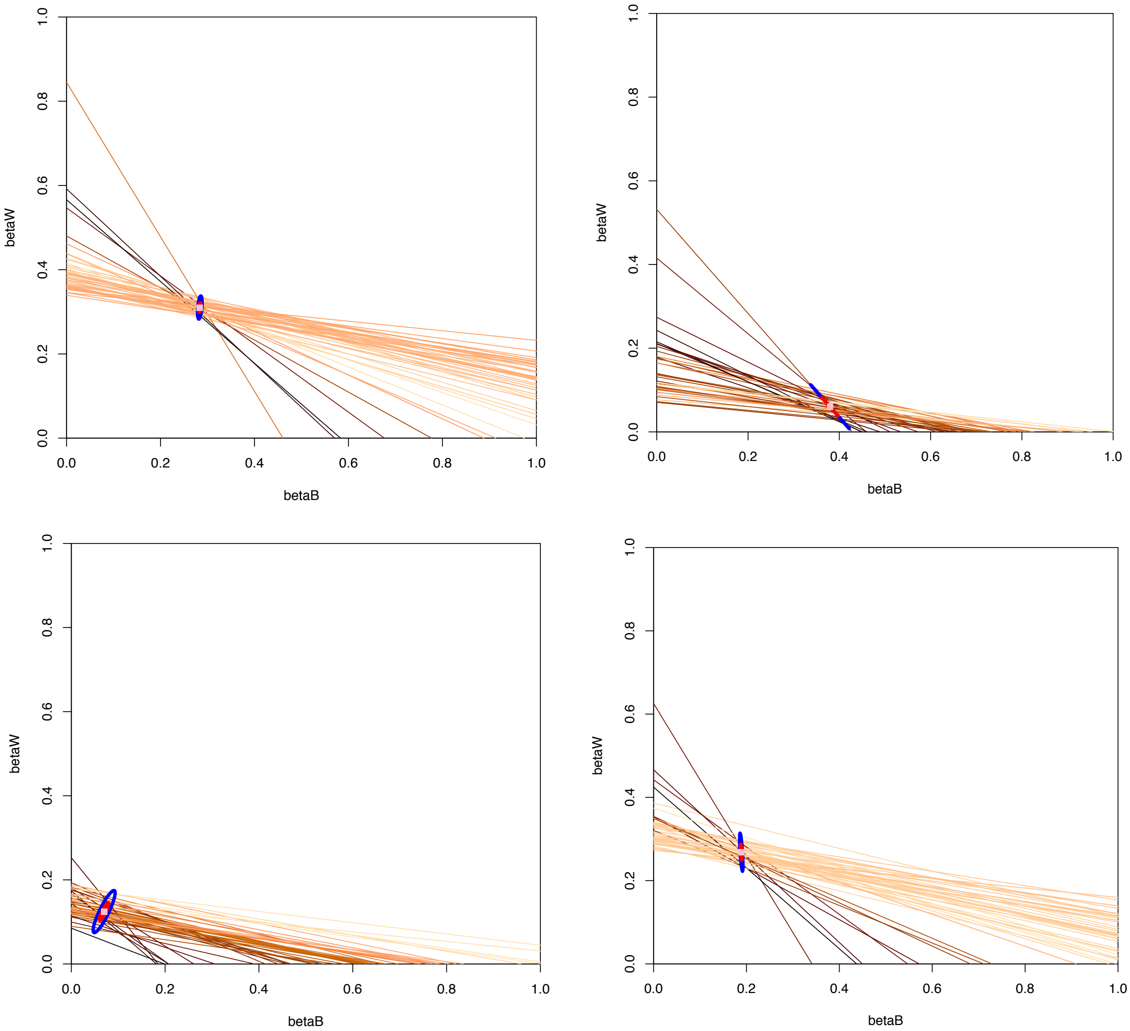


Figure 2: Less “Informative” Tomography Plots

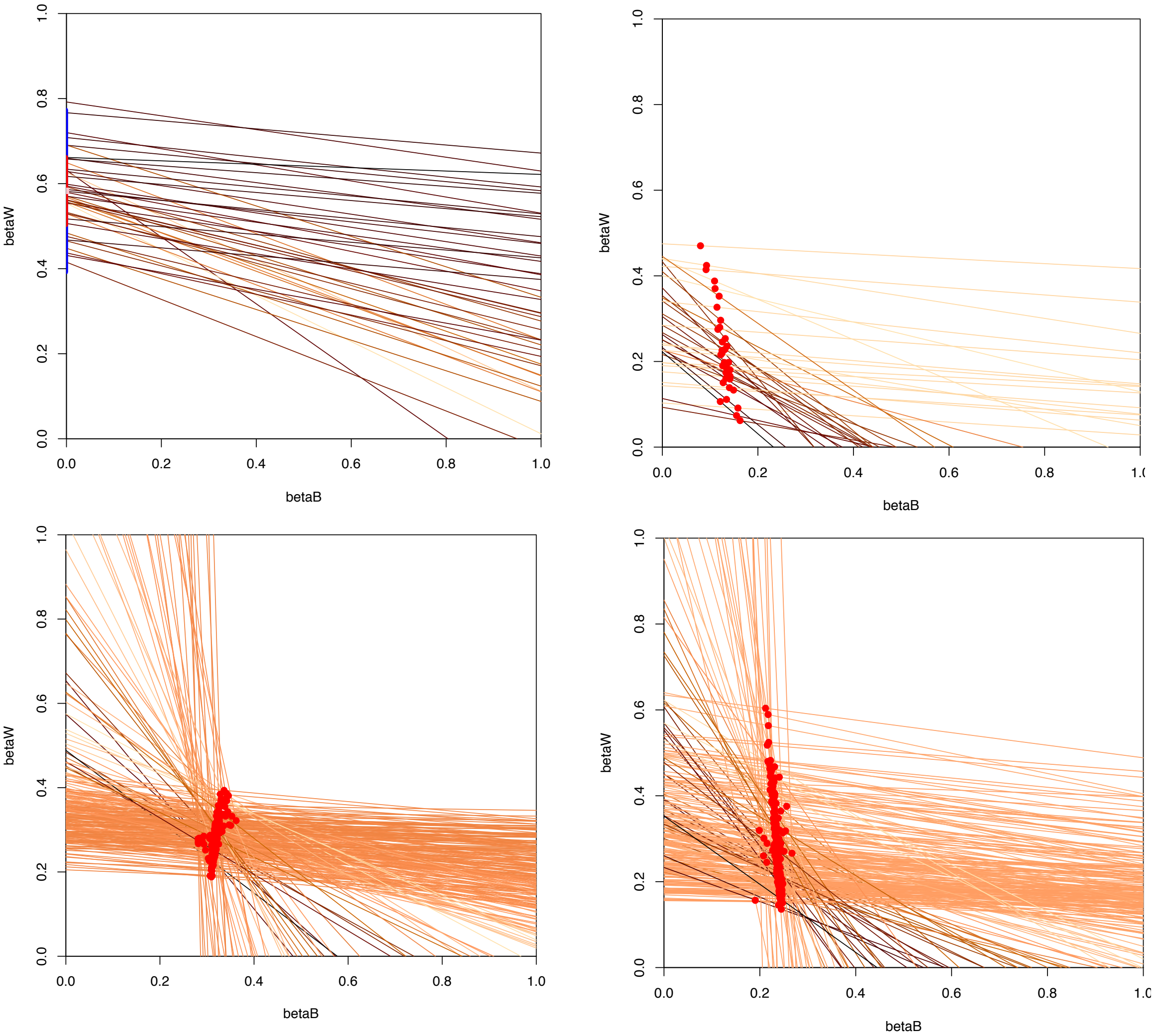


Figure 3

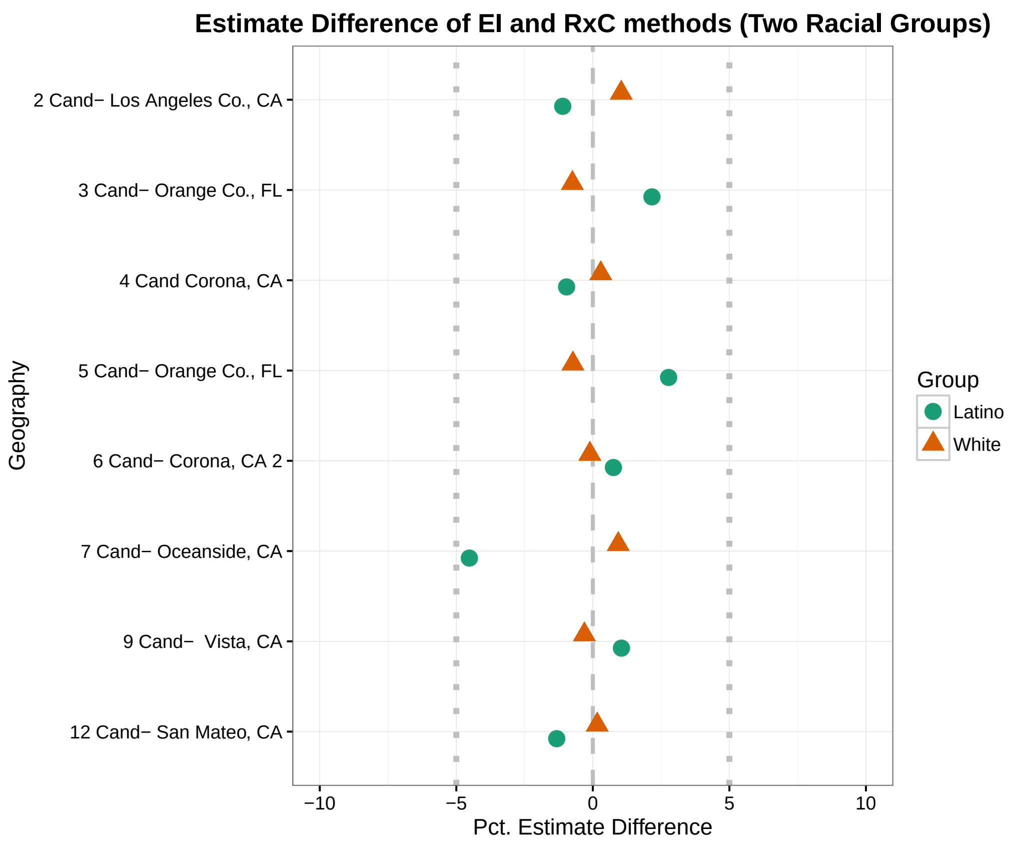

Estimate Difference of EI and RxC methods (Two Racial Groups)
Geography
2 Cand− Los Angeles Co., CA
3 Cand− Orange Co., FL
4 Cand Corona, CA
5 Cand− Orange Co., FL
6 Cand− Corona, CA 2
7 Cand− Oceanside, CA
9 Cand− Vista, CA
12 Cand− San Mateo, CA
−10
−5
0
5
10
Pct. Estimate Difference
Group
Latino
White

Figure 4

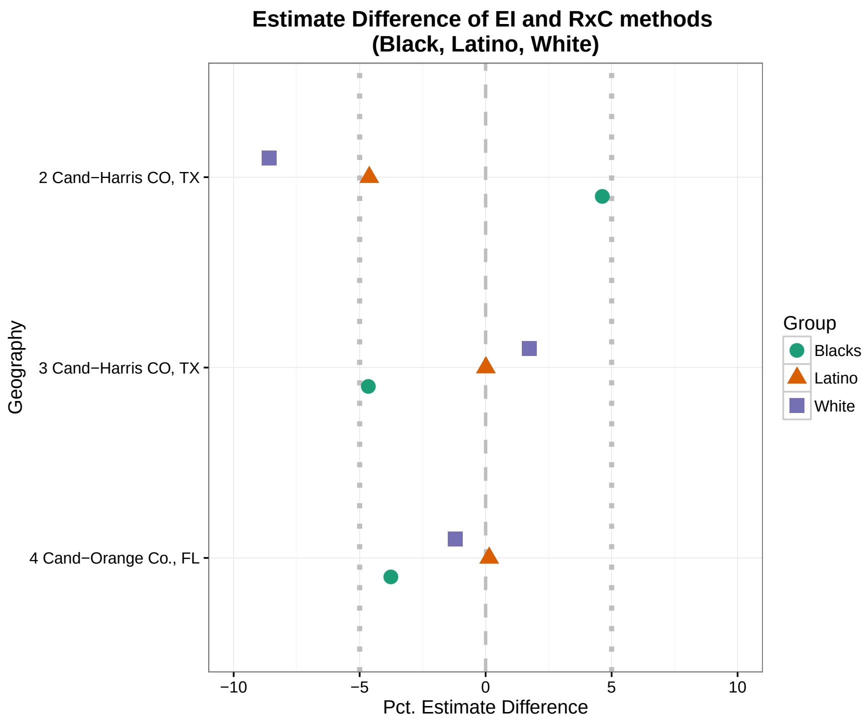

Estimate Difference of EI and RxC methods
(Black, Latino, White)
Geography
2 Cand−Harris CO, TX
3 Cand−Harris CO, TX
4 Cand−Orange Co., FL
−10
−5
0
5
10
Pct. Estimate Difference
Group
Blacks
Latino
White

Figure 5

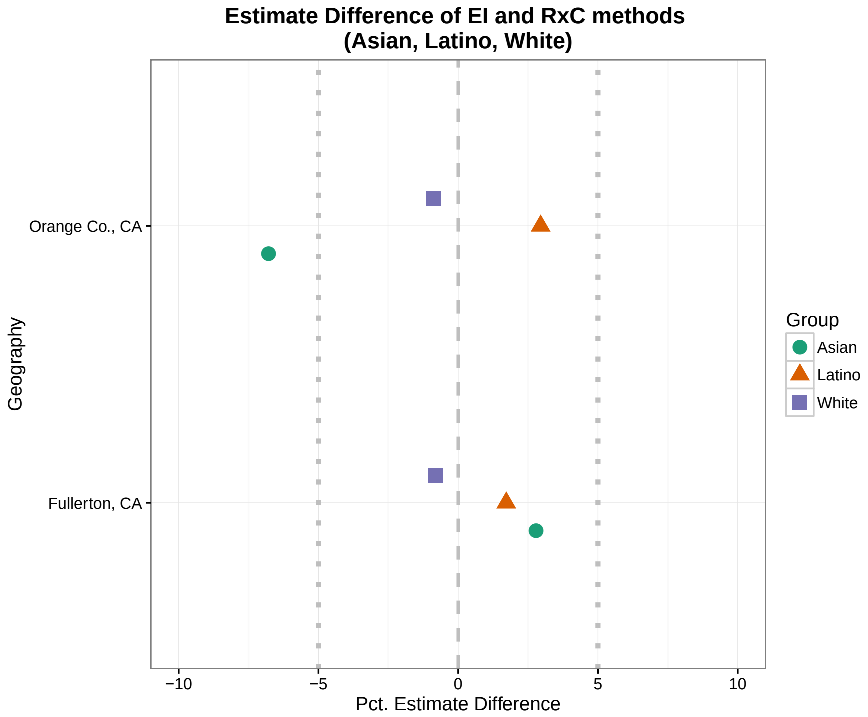
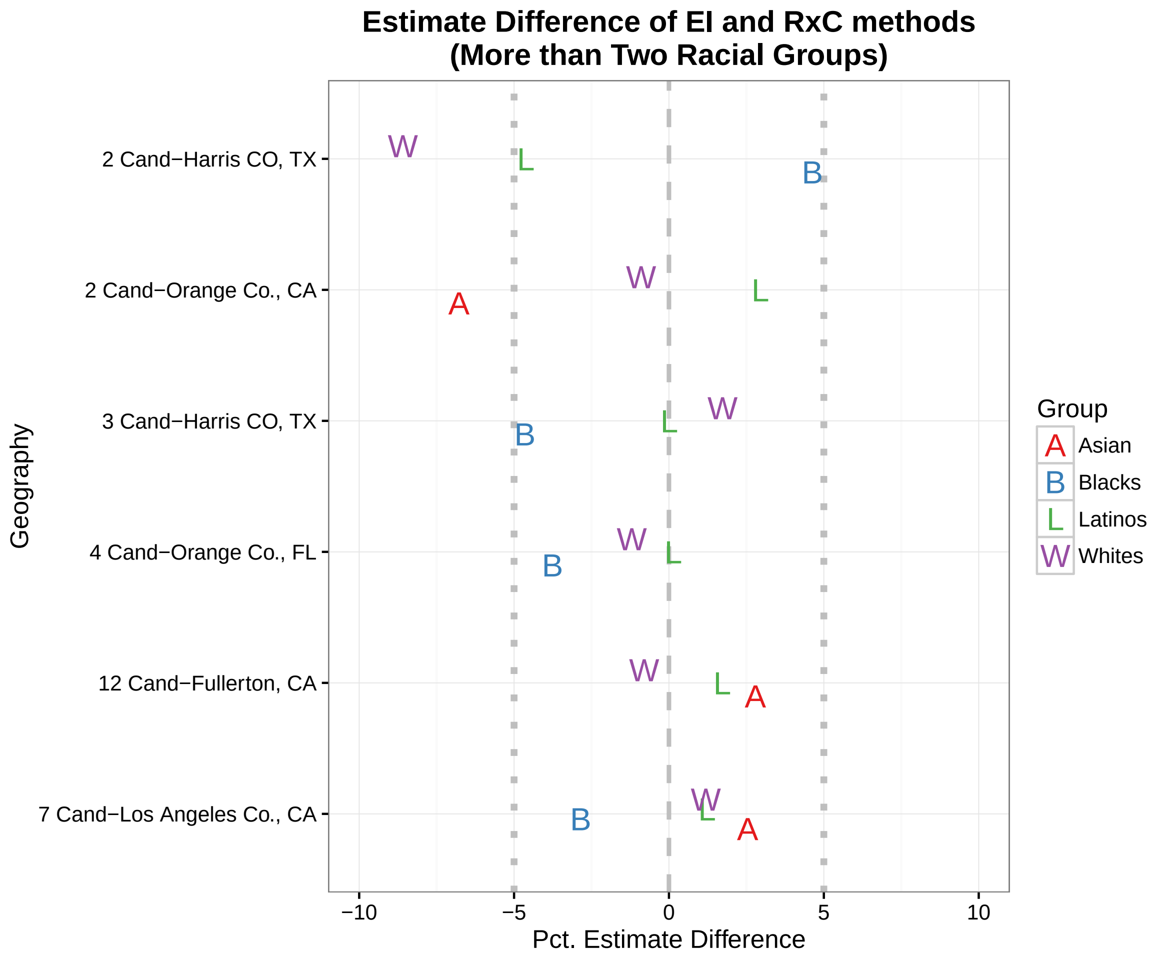
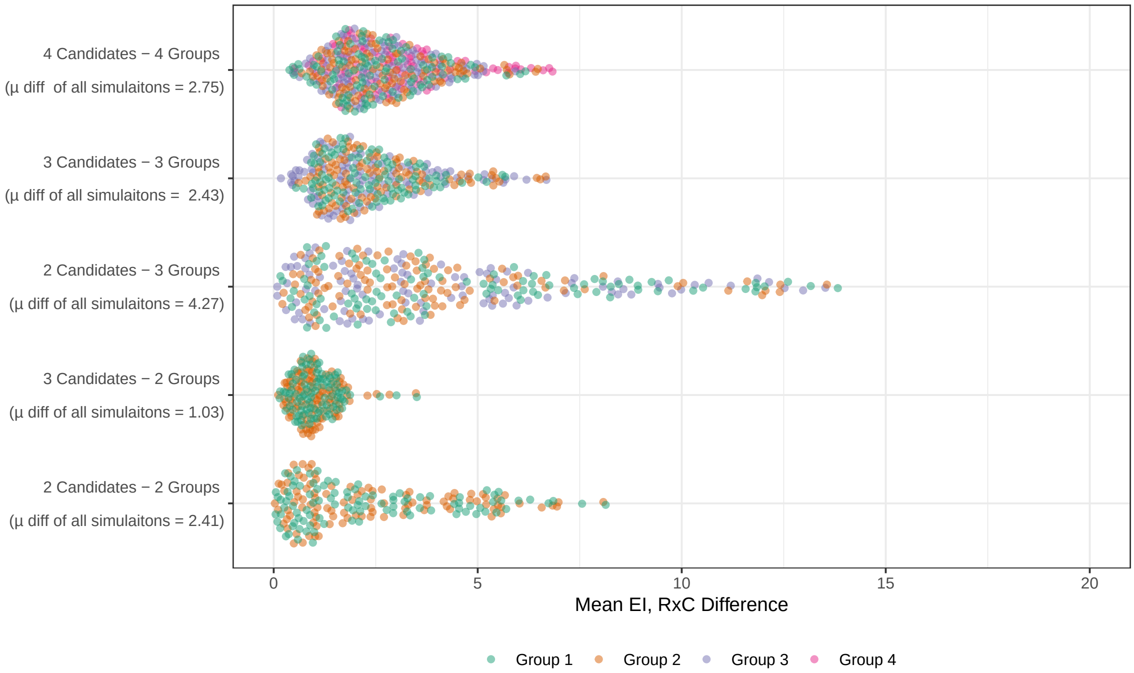

Estimate Difference of EI and RxC methods
(Asian, Latino, White)
Orange Co., CA
Fullerton, CA
Geography
−10
−5
0
5
10
Pct. Estimate Difference
Group
Asian
Latino
White

Figure 6



Estimate Difference of EI and RxC methods
(More than Two Racial Groups)
Geography
2 Cand−Harris CO, TX
2 Cand−Orange Co., CA
3 Cand−Harris CO, TX
4 Cand−Orange Co., FL
12 Cand−Fullerton, CA
7 Cand−Los Angeles Co., CA
−10
−5
0
5
10
Pct. Estimate Difference
Group
A Asian
B Blacks
L Latinos
W Whites

Figure 7: Simulated Data Results (500 datasets)



A Latino vs. Non-Latino

Table 11: Los Angeles County, CA Insurance Commissioner 2010 EI vs. EI:RxC Comparison

	Latino Vote			Non-Latino Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% De la Torre	84.11	82.94	-1.17	22.02	22.99	0.97
se	9.49	0.59		7.11	0.45	
% Jones	15.92	17.05	1.13	77.99	77.01	-0.98
se	9.51	0.58		7.11	0.45	
Total	100.03	100.00	-0.04	100.01	100.00	-0.01

Precinct n = 4980, Number of Candidates = 2

Table 12: Orange County, Florida School Board 2006 EI vs. EI:RxC Comparison

	Latino Vote			Non-Latino Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Flynn	0.8	3.6	2.8	57.8	57.7	-0.1
se	0.9	3.4		0.0	1.5	
% Kelly	15.7	18.7	3.0	32.4	30.5	-1.9
se	2.5	7.2		0.7	1.9	
% Cardona	94.3	96.5	2.2	8.1	7.4	-0.7
se	4.2	2.7		1.0	0.9	
Total	110.9	118.9	8.0	98.4	95.7	-2.7

Precinct n = 44, Number of Candidates = 3

Table 13: Corona, CA City Council 2006 EI vs. EI:RxC Comparison

	Latino Vote			Non-Latino Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Breitenbucher	19.6	18.1	-1.5	21.1	21.5	0.4
se	0.7	1.6		0.1	0.5	
% Montanez	35.9	34.9	-0.1	20.1	20.4	0.3
se	0.02	1.70		0.05	0.56	
% Spiegel	28.4	28.2	-0.2	30.9	31.0	0.1
se	0.6	1.1		0.3	0.3	
% Skipworth	18.8	18.6	-0.2	26.8	26.9	0.1
se	0.8	1.7		0.4	0.5	
Total	102.9	100.0	-2.9	99.1	100.0	0.9

Precinct n = 47, Number of Candidates = 4

Table 14: Orange County, Florida 2012 EI vs. EI:RxC Comparison

	Latino Vote			Non-Latino Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Clarke	24.7	23.6	-1.1	23.2	23.2	0.0
se	10.3	3.3		3.7	1.4	
% Damiani	10.7	15.5	4.8	37.2	35.3	-1.9
se	2.8	6.1		0.9	2.6	
% Lasso	13.3	12.2	-1.1	15.4	16.1	0.7
se	2.3	2.0		2.3	0.8	
% Aviles	35.2	38.0	2.8	2.7	2.0	-0.7
se	5.0	2.1		1.5	0.8	
% Pisano	12.0	11.0	-1.0	22.5	23.1	0.6
se	0.8	5.4		0.1	2.4	
Total	96.1	100.6	4.5	101.2	99.8	-1.4

Precinct n = 38, Number of Candidates = 5

Table 15: Corona, CA City Council 2004 EI vs. EI:RxC Comparison

	Latino Vote			Non-Latino Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Miller	20.7	15.9	-4.8	28.2	29.3	1.1
se	10.1	4.4		2.5	1.3	
% Melendez	38.5	39.2	0.7	4.5	4.4	-0.1
se	2.0	1.9		0.8	0.6	
% Nolan	18.6	16.3	-2.3	25.7	26.4	0.7
se	0.1	3.4		0.1	1.0	
% Humphrey	7.1	6.8	-0.3	12.4	12.5	0.1
se	1.7	2.3		0.4	0.6	
% Schnbal	2.5	3.0	0.5	8.5	8.6	0.1
se	2.1	1.1		0.7	0.3	
% Bennett	18.5	18.5	0.0	18.5	18.5	0.0
se	2.7	2.6		0.7	0.7	
Total	106.0	99.9	-6.1	97.9	100.0	2.1

Precinct n = 48, Number of Candidates = 6

Table 16: Oceanside, CA City Council 2012 EI vs. EI:RxC Comparison

	Latino Vote			Non-Latino Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Dykes	0.8	2.0	1.2	17.8	17.6	-0.2
se	0.8	1.7		0.0	0.6	
% Corso	9.4	15.8	6.4	20.8	21.9	1.1
se	3.8	3.7		0.4	0.8	
% Zerinik	8.3	9.1	0.8	6.7	6.5	-0.2
se	0.9	1.3		0.1	0.3	
% Snyder	6.8	6.6	-0.2	1.4	1.7	0.3
se	0.7	0.6		0.7	0.1	
% Sanchez	53.1	48.5	-4.6	21.8	22.7	0.9
se	8.2	4.5		2.0	1.0	
% Feller	7.6	10.7	3.1	25.5	24.7	-0.8
se	3.8	4.1		1.0	0.9	
% Knott	12.0	12.5	0.5	3.7	3.6	-0.1
se	1.0	1.1		0.2	0.2	
Total	98.2	105.6	7.4	98.0	98.9	0.9

Precinct n = 78, Number of Candidates = 7

Table 17: Vista, CA City Council 2012 EI vs. EI:RxC Comparison

	Latino Vote			Non-Latino Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% YoungRigby	9.0	8.7	-0.3	17.0	17.1	0.1
se	0.9	2.0		0.3	0.5	
% Miles	9.9	8.9	-1.0	3.0	3.2	0.2
se	1.5	1.4		0.3	0.3	
% Kaiser	2.5	2.8	0.3	18.5	18.3	-0.2
se	1.6	1.8		0.4	0.5	
% Campbell	15.0	14.9	-0.1	18.6	18.6	0.0
se	4.2	1.8		1.0	0.5	
% Lopez	37.9	38.9	1.0	6.0	5.6	-0.4
se	0.1	1.7		0.1	0.4	
% Garretson	2.7	2.6	-0.1	11.9	11.4	-0.5
se	2.3	2.1		0.1	0.7	
% Ford	7.5	2.3	-5.2	5.0	7.3	2.3
se	0.4	1.8		0.3	0.6	
% Staight	8.3	8.2	-0.1	3.3	3.4	0.1
se	1.5	1.2		0.2	0.3	
% Fleming	23.1	19.4	-3.7	13.0	13.2	0.2
se	8.3	3.4		2.0	0.9	
Total	116.2	107.0	-9.2	96.7	98.4	1.7

Precinct n = 36, Number of Candidates = 9

Table 18: San Mateo, CA 2010 Primary EI vs. EI:RxC Comparison

	Latino Vote			Non-Latino Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Gutierrez	32.8	27.7	-5.1	7.1	7.4	0.3
se	20.7	2.1		1.9	0.3	
% Lenning	5.3	1.6	-3.7	3.2	3.6	0.4
se	4.6	1.0		0.4	0.1	
% Martin	0.0	1.8	1.8	2.3	2.1	-0.2
se	0.0	0.7		0.0	0.1	
% McMicken	7.2	9.6	2.4	7.0	6.7	-0.3
se	4.7	1.4		0.4	0.2	
% Deligianni	2.6	2.8	0.2	4.9	4.9	0.0
se	2.2	1.2		0.2	0.1	
% Shiehk	0.9	0.4	-0.5	0.6	0.6	0.0
se	0.8	0.2		0.1	0.0	
% Nusbaum	1.2	4.1	2.9	3.2	3.4	0.2
se	1.1	1.0		0.5	0.1	
% Romero	43.1	41.8	-1.3	17.8	18.0	0.2
se	15.3	2.7		1.8	0.4	
% Blake	0.8	0.6	-0.2	5.7	5.9	0.2
se	0.8	0.5		0.1	0.2	
% Williams	0.1	2.5	2.4	1.6	1.4	-0.2
se	0.0	0.6		0.1	0.1	
% Torlakson	8.2	8.3	0.1	27.3	27.3	0.0
se	7.1	3.7		0.8	0.5	
% Aceves	1.4	5.3	3.9	17.9	17.5	-0.4
se	0.9	2.7		0.2	0.4	
Total	104.0	107.14	3.1	99.1	99.3	0.2

Precinct n = 433, Number of Candidates = 12

B Latino, Asian, & White

Table 19: Orange County, CA Insurance Commissioner 2010 EI vs. EI:RxC Comparison

	Latino Vote			Asian Vote			White Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Jones	11.9	8.9	-3.0	54.6	61.6	7.0	64.9	65.8	0.9
se	10.1	1.8		12.1	1.6		4.5	0.3	
% Delatorre	88.0	90.9	2.9	45.1	38.3	-6.8	35.0	34.1	-0.9
se	10.1	1.8		12.1	1.6		4.5	0.3	
Total	100.0	99.9	-0.1	99.7	100.0	0.3	99.9	100.0	0.1

Precinct n = 1941, Number of Candidates = 2

Table 20: Fullerton City, CA City Council 2012 EI vs. EI:RxC Comparison

	Latino Vote			Asian Vote			White Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Jaramillo	21.5	13.7	-7.8	1.9	1.9	0.0	4.3	6.3	2.0
se	4.2	2.6		1.6	1.6		2.0	0.8	
% Hakim	12.0	7.6	-4.4	7.6	3.0	-4.6	2.8	3.6	0.8
se	3.3	2.4		2.1	2.1		0.8	0.8	
% Alvarez	17.5	19.2	1.7	6.1	8.9	2.8	8.2	7.4	-0.8
se	6.1	1.6		4.0	1.9		2.5	0.5	
% Reid	4.7	4.8	0.1	1.9	1.5	-0.4	1.1	0.8	-0.3
se	1.3	0.4		1.1	0.5		0.5	0.1	
% Kiger	8.2	9.9	1.7	17.8	16.8	-1.0	11.5	10.7	-0.8
se	5.2	1.7		6.4	2.1		2.2	0.6	
% Levinson	1.7	1.6	-0.1	14.4	10.2	-4.2	7.2	7.3	0.1
se	1.4	0.9		9.8	1.3		1.5	0.3	
% Bartholomew	3.8	5.5	1.7	5.6	6.6	1.0	6.0	5.3	-0.7
se	2.8	1.0		4.1	1.2		1.2	0.3	
% Whitaker	12.2	13.6	1.4	20.1	20.9	0.8	13.0	12.7	-0.3
se	6.4	1.5		0.1	1.9		2.3	0.5	
% Bankhead	5.6	4.7	-0.9	7.1	6.4	-0.7	7.2	7.2	0.0
se	4.2	1.1		1.3	1.4		1.4	0.4	
% Flory	11.5	7.9	-3.6	6.5	4.8	-1.7	12.6	13.5	0.9
se	3.0	2.3		0.7	2.5		2.0	0.8	
% Rands	2.9	3.2	0.3	13.2	11.7	-1.5	8.2	8.4	0.2
se	2.3	1.5		4.0	2.0		2.0	0.5	
% Fitzgerald	7.7	7.6	-0.1	13.4	12.4	-1.0	14.9	15.4	0.5
se	5.8	2.2		2.6	2.7		1.7	0.7	
Total	109.8	99.9	-9.9	116.0	105.6	-10.4	97.5	99.2	1.7

Precinct n = 84, Number of Candidates = 12

C Latino, Black, & White

Table 21: Harris County, TX 2010 General EI vs. EI:RxC Comparison

	Latino Vote			Black Vote			White Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Uribe	73.7	69.1	-4.6	95.0	99.6	4.6	13.4	4.8	-8.6
se	11.2	1.0		4.4	0.3		8.1	1.1	
% Patterson	26.2	30.8	4.6	4.9	0.3	-4.6	86.6	95.1	8.5
se	11.2	0.9		4.4	0.3		8.1	1.1	
Total	100.0	100.0	0.0	99.9	100.00	0.1	100.0	99.9	-0.1

Precinct n = 885, Number of Candidates = 2

Table 22: Harris County, TX 2010 Primary EI vs. EI:RxC Comparison

	Latino Vote			Black Vote			White Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Earle	28.8	28.2	-0.6	45.7	50.5	4.8	53.0	53.4	0.4
se	13.1	1.1		7.4	1.3		11.0	1.2	
% Katz	7.5	12.7	5.2	7.6	12.0	4.5	17.3	15.0	-2.3
se	2.7	0.7		3.0	0.9		6.0	0.8	
% Chavez	59.0	59.0	0.0	42.0	37.4	-4.6	29.7	31.5	1.8
se	12.9	1.1		0.0	1.4		10.5	1.2	
Total	95.4	99.9	4.5	95.3	99.9	4.7	100.1	100.0	-0.1

Precinct n = 885, Number of Candidates = 3

Table 23: Orange County, FL 2008 Soil/Water Board EI vs. EI:RxC Comparison

	Latino Vote			Black Vote			White Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Cardona	65.7	65.8	0.1	24.2	20.5	-3.7	17.7	16.5	-1.2
se	6.0	0.9		2.7	0.6		3.0	0.3	
% Hamada	19.3	20.1	0.8	32.6	34.5	1.9	29.9	31.0	1.1
se	1.4	1.1		2.2	0.8		0.5	0.4	
% Whiting	2.7	2.5	-0.2	14.2	14.8	0.6	26.0	27.0	1.0
se	2.1	0.9		1.5	0.7		2.2	0.3	
% Hamilton	11.0	11.4	0.4	29.0	30.0	1.0	24.5	25.4	0.9
se	2.6	0.9		0.6	0.6		2.6	0.3	
Total	98.8	100.0	1.2	100.2	99.9	-0.3	98.2	100.0	1.8

Precinct n = 252, Number of Candidates = 4

D Latino, Black, Asian, & White

Table 24: Los Angeles, CA 2010 State Attorney (General) EI vs. EI:RxC Comparison

	Latino Vote			Black Vote			Asian Vote			White Vote		
Candidate	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff	**EI**	**RxC**	Diff
% Harris	8.8	4.2	-4.6	63.0	72.7	9.7	17.3	20.5	3.2	33.9	33.9	0.0
se	6.1	0.4		18.9	0.6		14.8	0.8		9.6	0.3	
% Delgadillo	39.2	40.4	1.2	12.3	9.4	-2.9	11.8	14.4	2.5	10.9	12.1	1.2
se	8.0	0.3		4.2	0.4		4.2	0.6		5.8	0.2	
% Lieu	4.3	3.7	-0.6	0.8	0.8	0.0	28.7	26.8	-1.9	17.5	15.2	-2.3
se	3.7	0.3		0.7	0.5		13.2	0.8		8.0	0.3	
% Kelly	9.6	11.5	1.9	5.9	8.6	2.7	8.9	14.8	5.9	17.7	16.8	-0.9
se	3.2	0.2		3.1	0.4		4.4	0.5		2.8	0.2	
% Torrico	17.8	18.2	0.4	4.6	3.7	-0.9	6.4	11.3	4.9	9.8	10.1	0.3
se	4.7	0.2		2.5	0.3		3.5	0.5		4.5	0.2	
% Nava	16.4	17.7	1.3	3.5	1.1	-2.4	5.4	8.1	2.7	6.5	6.9	0.4
se	5.2	0.2		2.1	0.3		2.6	0.4		3.8	0.1	
% Schmier	4.0	4.0	0.0	1.8	3.5	1.7	3.4	3.8	0.4	5.0	4.6	-0.4
se	3.1	0.1		1.6	0.2		0.7	0.3		2.9	0.1	
Total	100.4	99.9	-0.5	92.1	100.1	8.0	82.2	100.0	17.8	101.6	99.9	-1.7

Precinct n = 4974, Number of Candidates = 7

E Congruence Comparison

The full vote choice results of these elections can be found in Tables 11 – 24 but in this section we detail how congruent the ecological estimates are across the EI and RxC models. We should note that as more candidates are added the process becomes more complex, but the same underlying principles highlighted in the 2x2 case apply.

Tables 25 – 28 show that across different types of elections – some with more candidates and some with more groups of voters – the iterative EI and simultaneous RxC are very congruent. The model congruence scores capture the relative rank of the candidates, the size of the gap between first and second choice, and the size of the gap between minorities and Anglos. In the examples provided below, we find very strong evidence of model congruence across different election settings.

Table 25: 2x4 Congruence table for Corona, CA City Council 2006

	EI	RxC	Congruence
MV1-WV for MC1	15.8	14.5	0.9142
MC1 preferred by MV1	Yes	Yes	1
MC1 pref rate	3.8	3.4	0.8873
MC1 blocked by WV	Yes	Yes	1
MC1 block rate	-5.4	-5.3	0.9813
MC model congruence score			0.9566
MV1-WV for WC1	-2.5	-2.8	0.8868
WC1 preferred by WV	Yes	Yes	1
WC1 pref rate	2.05	2.1	0.9759
WC1 blocked by MV1	Yes	Yes	0.5
WC1 block rate	-3.75	-3.4	0.8873
WC model congruence score			0.85
Total model congruence score			0.9033

Note: see Table 13 for actual polarized voting results for EI and RxC

Table 26: 2x5 Congruence table for Orange County, FL Commissioner 2012

	EI	RxC	Congruence
MV1-WV for MC1	32.5	36	0.8978
MC1 preferred by MV1	Yes	Yes	1
MC1 pref rate	5.3	7.2	0.6867
MC1 blocked by WV	Yes	Yes	1
MC1 block rate	-17.3	-16.7	0.9646
MC model congruence score			0.9098
MV1-WV for WC1	-26.5	-19.8	0.7106
WC1 preferred by WV	Yes	Yes	1
WC1 pref rate	7	6.1	0.8544
WC1 blocked by MV1	Yes	Yes	0.8
WC1 block rate	-12.25	-11.3	0.9149
WC model congruence score			0.856
Total model congruence score			0.8829

Note: see Table 14 for actual polarized voting results for EI and RxC

Table 27: 3x2 Congruence table for Harris County, TX Lt. Gov Dem Primary 2010

	EI	RxC	Congruence
MV1 - Latinos			
MV1-WV for MC1	60.3	64.3	0.9358
MC1 preferred by MV1	Yes	Yes	1
MC1 pref rate	23.8	19.2	0.7855
MC1 blocked by WV	Yes	Yes	1
MC1 block rate	-36.6	-45.2	0.7908
MC1 model congruence score			0.9024
WV - Whites			
MV1-WV for WC1	-60.4	-64.3	0.9374
WC1 preferred by WV	Yes	Yes	1
WC1 pref rate	36.6	45.2	0.7908
WC1 blocked by MV1	Yes	Yes	1
WC1 block rate by MV1	-23.8	-19.2	0.7855
WC model congruence score			0.9028
MV2 - Blacks			
MV2-WV for MC1	81.6	94.8	0.8503
MC1 preferred by MV2	Yes	Yes	1
MC1 pref rate	45.1	49.7	0.9029
MC1 blocked by WV	Yes	Yes	1
MC1 block rate	-36.6	-45.2	0.7908
MC1 model congruence score			0.9088
WV - Whites			
MV2-WV for WC1	-81.7	-94.8	0.8516
WC1 blocked by MV2	Yes	Yes	1
WC1 block rate by MV2	-45.1	-49.7	0.9029
WC model congruence score			0.9182
Total model congruence score			0.9081

Note: see Table 22 for actual polarized voting results for EI and RxC

Table 28: 4x7 Congruence Table for Los Angeles County, CA Primary election for Attorney General 2010

	EI	RxC	Congruence
MV1 - Latinos			
MV1-WV for MC1	28.3	28.3	1
MC1 preferred by MV1	Yes	Yes	1
MC1 pref rate	10.7	11.1	0.9633
MC1 blocked by WV	Yes	Yes	1
MC1 block rate	-11.5	-10.9	0.9464
MC1 model congruence score			0.9819
WV - Whites			
MV1-WV for WC1	-25.1	-29.7	0.8321
WC1 preferred by WV	Yes	Yes	1
WC1 pref rate	8.1	8.6	0.9459
WC1 blocked by MV1	Yes	Yes	1
WC1 block rate by MV1	-15.2	-18.1	0.8258
WC model congruence score			0.9208
MV2 - Blacks			
MV2-WV for MC2	29.1	38.8	0.7143
MC2 preferred by MV2	Yes	Yes	1
MC2 pref rate	25.4	31.7	0.7789
MC2 blocked by WV	No	No	1
MC2 block rate*	8.1	8.6	0.9459
MC2 model congruence score			0.8878
WV - Whites			
MV2-WV for WC1	-29.1	-38.8	0.7143
WC1 blocked by MV2	No	No	1
WC1 block rate by MV2	25.4	31.7	0.7789
WC model congruence score			0.8311
MV3 - Asians			
MV3-WV for MC3	11.2	11.6	0.9649
MC3 preferred by MV3	Yes	Yes	1
MC3 pref rate	5.7	3.2	0.4237
MC3 blocked by WV	Yes	Yes	1
MC3 block rate	-8.2	-9.4	0.8689
MC3 model congruence score			0.8515
WV - Whites			
MV3-WV for WC1	-16.6	-13.4	0.7867
WC1 blocked by MV3	Yes	Yes	1
WC1 block rate by MV3	-5.7	-3.2	0.4237
WC model congruence score			0.7368
Total model congruence score			0.8717

Note: see Table 24 for actual polarized voting results for EI and RxC