

The Geography of Racially Polarized Voting: Calibrating Surveys at the District Level*

Shiro Kuriwaki[†] Stephen Ansolabehere[‡] Angelo Dagonel[§] Soichiro Yamauchi[¶]

July 2023

American Political Science Review

DOI: [10.1017/S0003055423000436](https://doi.org/10.1017/S0003055423000436)

Abstract

Debates over racial voting, and over policies to combat vote dilution, turn on the extent to which groups' voting preferences differ and vary across geography. We present the first study of racial voting patterns in every congressional district in the US. Using large-sample surveys combined with aggregate demographic and election data, we find that national-level differences across racial groups explain 60 percent of the variation in district-level voting patterns, while geography explains 30 percent. Black voters consistently choose Democratic candidates across districts, while Hispanic and White voters' preferences vary considerably across geography. Districts with the highest racial polarization are concentrated in the parts of the South and Midwest. Importantly, multi-racial coalitions have become the norm: in most congressional districts, the winning majority requires support from non-White voters. In arriving at these conclusions, we make methodological innovations that improve the precision and accuracy when modeling sparse survey data.

*We thank Matt Barreto, Andrew Gelman, Yair Ghitza, Rob Griffin, Andrew Hall, Kabir Khanna, Lauren Kennedy, Mayya Komisarchik, Jeff Lewis, Maxwell Palmer, Douglas Rivers, Jonathan Robinson, Jonathan Rodden, Yajuan Si, members of the Democracy and Polarization Lab at Stanford, and participants of Polmeth 2022 for their valuable suggestions. Data and software to implement our methods are publicly available through the R packages [ccesMRPprep](#), [ccesMRPrun](#), and [synthjoint](#).

[†]Corresponding Author. Assistant Professor of Political Science, Department of Political Science and Institution for Social and Policy Studies, Yale University, United States. Website: www.shirokuriwaki.com. Email: shiro.kuriwaki@yale.edu. ORCID: [0000-0002-5687-2647](https://orcid.org/0000-0002-5687-2647).

[‡]Frank G. Thomson Professor of Government, Department of Government, Harvard University, United States. Website: scholar.harvard.edu/sansolabehere. Email: sda@gov.harvard.edu.

[§]PhD candidate, Department of Government, Harvard University, United States. Website: sites.google.com/g.harvard.edu/dagonel. Email: dagonel@g.harvard.edu.

[¶]Independent scholar. Website: soichiroy.github.io. Email: soichiro@alumni.princeton.edu. ORCID: [0000-0003-0554-2717](https://orcid.org/0000-0003-0554-2717).

Table of Contents

Introduction	3
Measuring Vote Choice By Race Using Surveys	6
Quantities of Interest	7
Survey Data	7
Challenges in Survey Inference	9
Calibrating Surveys at the District Level	10
Hierarchical Regression	11
Survey-assisted Synthetic Target Estimation	13
Two-way Calibration to Election Results and External Surveys	15
Evaluation and Validation	17
The Political Geography of Vote Choice by Race	21
The Racial Gap at the District Level	21
Group Cohesion	24
Groups and Coalitions	32
Conclusion	35
References	40
A Modeling Methodology	A1
A.1 Notation	A1
A.2 About the Cooperative Congressional Election Study	A2
A.3 Implementation of Hierarchical Regression	A3
A.4 Connections with Ecological Inference Estimators	A5
A.5 Population Statistics for Target Estimation	A9
A.6 Survey-assisted Synthetic Target Estimation	A9
A.7 Estimation of Joint Demographic Distributions	A13
A.8 Estimation of a Calibrated Turnout Model	A14
A.9 Grouping of States	A15
A.10 How Modeling Choices Affect Estimates	A15
B Additional Analyses	A17
B.1 Asian Americans	A17
B.2 Comparison with Other Surveys	A18
B.3 Details on Florida Validation Analysis	A19
B.4 Comparison with Ecological Inference in 50 States	A20
C Additional Estimates and Standard Errors	A24
C.1 Additional Estimates	A24
C.2 2020 Estimates	A25
C.3 Racial Gap Estimates	A25
C.4 Standard Errors	A29

INTRODUCTION

The United States is a nation divided by color. Generations of social science have documented that the areas of the country with high Black and Hispanic populations, such as the Black Belt counties of the South or the Hispanic neighborhoods in southwestern cities, usually vote for candidates different than those who win in predominantly White areas (Key 1948; Kousser 1974; Davidson and Grofman 1994). Most scholarship attributes such differences to strong group identities and even animus between White and non-White groups (Grofman 1991; Dawson 1995; King 1996; Kinder and Sanders 1996). Yet other scholars identify instances where context and geography generate variation in group behavior (Gay 2001; Hopkins 2010; Enos 2017; Gimpel et al. 2020). It is difficult to gauge the relative importance of racial groups and of geography from aggregate data (Gelman 2009). The patterns that scholars think they see in aggregate election outcomes could, as Freedman et al. (1991) argued, be due to differences across neighborhoods instead of differences between groups.

This study uses survey data to offer the first nationwide assessment of racial voting patterns at the congressional district (CD) level. We find that racial groups on average vote differently from each other, but the differences across Black, Hispanic, White, and Other voters also vary across states and across CDs within states. Specifically, 60 percent of the variation in 2016 Presidential vote shares at the district-racial group level is explained by national differences across racial groups. Variation across regions, states, and CDs explains another 30 percent of the variation, with CDs explaining twice as much variation as states do. Nearly everywhere, Black voters vote at high rates for Democratic candidates. The voting behavior of Hispanic voters and of White voters is far more variable across geography. At least as far as racial group voting is concerned, then, our results reveal that national divisions (Hopkins 2018; Kinder and Sanders 1996) are more prominent than local divisions (Cramer 2016; Gimpel et al. 2020).

One of the most striking features of contemporary politics that emerge is the degree to which both the Democratic and Republican parties rely on the support of White and minority voters. In most CDs the winning candidate relied on some support from at least one minority group in order to win a majority of votes. In two-thirds of Democratic-leaning districts, the Democratic majority consisted of combinations of Black, Hispanic, and White voters, with no single group determining the majority. In one out of five Republican-leaning districts, the White vote alone was not sufficient to reach a majority in 2016. Republican majorities in those districts depended on the Republican votes of Hispanic or other racial minorities.

Our study makes contributions in three fields: race and voting behavior, survey methodology, and election law. First, we contribute to the study of racial voting in the United States by providing the first comprehensive picture of racial group voting patterns across all congressional districts. This allows us to test competing claims about whether racial group differences in vote choice are explained by national groups or to the local context of where people live. Survey research has long documented national group differences: The 2016 Exit Polls estimates that about 60 percent of White voters nationwide voted for the Republican Presidential candidate, while 92 percent of Black voters, 70 percent of Hispanic voters, and 70 percent of Asian voters voted for the Democrat. A lengthy literature argues that people have clear, distinct vote preferences that are rooted in racial group identities (e.g., Kinder and Sanders 1996). Increasingly, though, scholarship on race and American elections has emphasized the importance of geographic variation in voting behavior of racial groups (Cho 1995; Gelman 2009; Erikson 2010; Enos 2017), especially differences between urban and rural areas (McKee 2008; Rodden 2019), differences across counties (Acharya, Blackwell, and Sen 2018), and variation across districts (Donovan 2010). We show that racial groups are not monolithic in their voting behaviors (Cho 1995; Erikson 2010). We find that Black voters vote overwhelmingly Democratic regardless of where they live, but the voting preferences of White and Hispanic voters vary across regions, states, and, even, districts within

states. That said, we show that the variation in district-level Presidential voting behavior is explained primarily by differences across racial groups at the national-level. In short, who people are is more important than where they live.

Second, we improve upon existing survey modeling methods by developing two new calibration techniques. Even large surveys such as exit polls suffer from insufficient sample sizes to estimate group behavior in each congressional district with much precision. Political scientists have recently deployed Multilevel Regression and Poststratification (MRP) for small area estimation. MRP uses hierarchical modeling and calibration weights derived from population statistics to make subgroup estimates more precise (Gelman and Little 1997). However, the characteristics of the target population required for weighting are often not available (Lee-mann and Wasserfallen 2017). This is partly why MRP estimates for political polls can still be unrepresentative of small areas (Lauderdale et al. 2020). We create demographic weighting targets of the electorate at the district level using a technical innovation we call survey-assisted synthetic target estimation. This improves weighting estimates to the distribution of race, age, sex, and education within the actual electorate of each district. Furthermore, we develop a two-way survey calibration, which simultaneously calibrates estimates to both election results by geography and an external survey, instead of only to geography (e.g., Ghitza and Gelman 2013; Rosenman, McCartan, and Olivella 2023). This adjustment reduces bias in the estimates of group voting behavior at the CD-level due to unobservable selection bias in the survey.

Third, our analysis informs on-going debates over election laws and voting rights. Measuring the extent of racially polarized voting has been social science’s key contribution to the development of election laws designed to protect minority voters (Pildes 2001). Significant questions remain as to where and how the law ought to apply. If voting patterns largely reflect national group differences, then a broad national policy is the best approach to fight vote dilution, as some have argued (Charles and Fuentes-Rohwer 2014; Cain and Zhang 2016).

If a racial group's preferences vary substantially across geography, then a narrowly tailored approach to the law may be more effective and appropriate, as in the case of *Alabama Legislative Black Caucus v. Alabama* (2015). Almost all of the prior scholarship stops at state-level estimates (Shaw 1997; Ansolabehere, Persily, and Stewart 2009, 2012; Stephanopoulos 2016; Smith, Kreitzer, and Suo 2020). Only Elmendorf and Spencer (2015) and Ghitza and Gelman (2020) provide estimates at the county-level, but those analyses do not apply to the most common application – redistricting. Our findings regarding the substantial national division in racial voting patterns and the variations across areas point to the need for both a broad national approach, as well as some degree of narrow tailoring to particular districts and areas.

MEASURING VOTE CHOICE BY RACE USING SURVEYS

Social scientists frequently seek to measure how individual behavior varies across geographic areas and demographic subgroups. Aggregate data alone cannot distinguish how much of the variation is attributable to differences across demographic groups or variation across levels of geography. Ecological inference techniques attempt to infer the characteristics of groups by studying how aggregate outcomes vary by the composition of the group in each aggregate unit. In doing so, standard ecological inference assumes that variations in the behavior of a racial group in different geographies are independent of the composition of the group (Freedman et al. 1991; Cho 1998; Ansolabehere and Rivers 1995, see also Online Appendix A.4 for a formal exposition). Moreover, ecological inference struggles to identify any estimate at all in districts where racial groups are highly integrated within all precincts. Our study estimates the contributions of racial groups and geography directly using individual-level survey data, rather than make assumptions about the structure of aggregate data.

Quantities of Interest

Our central quantity of interest is the share of a candidate's support among a particular racial group in a particular district. Let $i \in \{1, \dots, N\}$ index all individual voters who voted for a Republican or Democratic Presidential candidate. We index the voter's racial group as $g \in \{1, \dots, G\}$, and congressional districts (CDs) as $j \in \{1, \dots, J\}$ with $J = 435$. We let \mathcal{I}_{gj} denote the set of voters with race g in district j , and N_{gj} denote the total number of such voters. A voter has $Y_i = 1$ if they voted for the Republican candidate, and $Y_i = 0$ if they voted for the Democratic candidate. We are interested in the Republican voteshare in each racial group in each congressional district,

$$\tau_{gj} = \frac{1}{N_{gj}} \sum_{i \in \mathcal{I}_{gj}} Y_i,$$

To report racial group differences, we then define the racial gap in district j between Whites and a non-White racial group g as

$$\text{Racial Gap}_{gj} = \tau_{\text{white}, j} - \tau_{gj}. \quad (1)$$

Voting is racially polarized if the vote shares of white and non-white groups are on the opposite sides of 50 percent.

Survey Data

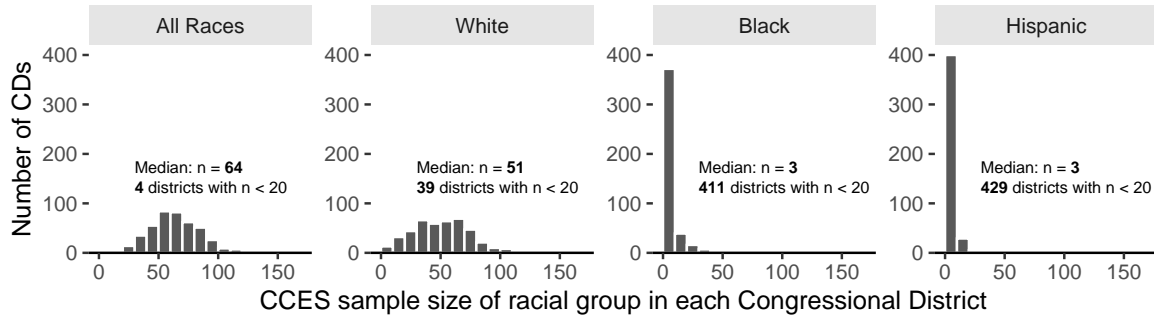
We use the 2016 Cooperative Congressional Election Study and 2020 Cooperative Election Study. These surveys interview over 60,000 respondents nationwide, and validate records by matching them to vote histories available from voter registration data. It is one of the largest political surveys of American Politics and contains the information necessary for our model: vote choice, congressional district, turnout matched from the voter file, racial group

identification, and other demographics.

Throughout this study, we group race into four categories: $g \in \{\text{White, Black, Hispanic, All Others}\}$. Black is used interchangeably with African American, and Hispanic is used interchangeably with Latino in our analysis. As with the US Census, the CCES asks Hispanic ethnicity as a separate question from race. Respondents who are “any-part” Hispanic are coded as Hispanic (see Online Appendix A.2 for more detail of the survey and question wording). In the 2016 CCES, about 2 percent of White respondents and about 2 percent of Black respondents also identify as Hispanic. Coding these respondents as Hispanics rather than Whites or Blacks typically increases the estimated Republican vote among the Hispanic group by 2 percentage points and decreases the estimated Republican vote of Whites and Blacks by a few tenths of a percentage point. Separating voters in the “Other” categories into Asian, Native American, and multi-racial respondents requires even more data than is available with the CCES and Census surveys. We replicate our analysis for Asian American voters limited to four states where they comprise over 10 percent of the population (Online Appendix B).

Our outcome of interest is the respondent’s self-reported vote in the election for President in each year. We focus on the office of President in this study because it is the only ticket elected nationwide, and our goal is a 50-state, 435-district comparison of voter preferences between the same set of contesting candidates. Party choice across votes for U.S. Senate, U.S. House, and Governor within a single respondent may differ, though less so in the modern era (Jacobson 2015; Kuriwaki 2021). We limit our analysis to validated voters, who are respondents whose personally identifiable information records have been matched to public voter rolls. We also limit our analysis to those who reported voting for a major party Presidential candidate. The choice to drop third-party voters is in part for the convenience of modeling the outcome as a binary variable, but also for comparability across states and districts. Working with voteshares as a proportion of the two-party vote is also desirable because the two-party

Figure 1 – Sample sizes in each congressional district in the 2016 Cooperative Congressional Election Study. Samples are limited to validated voters who report voting for Donald Trump or Hillary Clinton, which we use in the survey model.



vote decides who wins the district or state. Thus, we focus on a subset of $n = 28,462$ of the entire CCES in 2016 and $n = 34,539$ for 2020.

Challenges in Survey Inference

Even with large surveys, including exit polls and the CCES, the number of observations for sub-state level geographies is still too small to estimate with precision individual-level variables within districts. In the subsample of the 2016 CCES studied here, the median sample size validated to be voters per congressional district is $n = 64$, and the median district only includes $n = 13$ voting respondents who are non-White (Figure 1). Recently, researchers have studied similarly sized samples in counties to draw inferences about racial groups' preferences in these local geographies (Acharya, Blackwell, and Sen 2018; Kuziemko and Washington 2018). Inferences from small subgroups can be dangerously misleading (Ansolabehere, Luks, and Schaffner 2015). Even with a sample of $n = 50$ in each group, the margin of error in the difference in proportions of those two groups is roughly 20 percentage points with a simple random sample, making any comparison uninformative. Hierarchical modeling proposed here offers a potential solution.

Another challenge for survey data is that the sample of each race and district subgroup

may be unrepresentative of the population counterpart (Grimmer et al. 2018). The CCES release includes poststratification weights that render its state subsamples representative of each state. However, these weights have no guarantees of making *district* subsamples representative of each district, and may only increase variance while failing to eliminate selection bias (Kuriwaki 2021, ch.4).

Some summary statistics of the population are known. A congressional district’s composition of race and other demographics of adult citizens are reported by high-quality Census surveys. The total presidential vote in each district, across all races, is also known from election results. We model the noisy and potentially unrepresentative survey data to improve the estimate of the quantity of interest τ_{gj} by calibrating to Census statistics and electoral results.

CALIBRATING SURVEYS AT THE DISTRICT LEVEL

Estimation proceeds in three steps: (1) conduct a hierarchical regression to predict the probability of a Republican vote in each geographic and demographic subgroup, (2) weight these probabilities to the estimated population size of these subgroups, and (3) calibrate these estimates to external targets at higher levels of aggregation. Steps (1) and (2) together constitute MRP (Gelman and Little 1997). Dozens of studies, including many that use the CCES, have implemented MRP to estimate subgroup political behavior at the national or state levels (e.g., Lax and Phillips 2009; Warshaw and Rodden 2012; Broockman and Skovron 2018; Elmen-dorf and Spencer 2015; Hertel-Fernandez, Mildemberger, and Stokes 2019).

We make two innovations to existing MPR methods. First, we develop in step (2) a method of *survey-assisted synthetic target estimation* that combines partially available Census statistics into a single joint distribution. Post-stratification in prior MRP work has been limited by the lack of population variables jointly available in off-the-shelf datasets. Using our method, we can guarantee through weighting that the estimates are representative of the age group,

sex, race, and education level of the population (according to the ACS) and the turnout rate in every congressional district. But this weighting may still be incomplete because of unmeasured selection bias. Second, we develop in step (3) a method of *two-way calibration* that matches targets both by geography and racial group, further reducing the potential bias in our estimates.

We generate estimates for both the 2016 CCES and the 2020 CES. Because there are no substantive differences in the methodology, we use the 2016 data as the running example and present 2020 results in Online Appendix C. In addition, we offer a validation analysis of voter registration by party and race in the state of Florida to show the accuracy and coverage of our approach compared to only using precinct-level aggregate data (ecological inference).

Hierarchical Regression

Traditional survey weights assign weights to each subgroup of respondents based on the demographic composition of the population within specific strata (such as states, regions, or districts). The estimated average vote Y among subgroups of respondents may be highly imprecise because subgroup samples within strata become very small. MRP compensates for the sparsity of data within strata samples by borrowing information across strata about particular groups. We estimate a hierarchical regression that returns the probability that voters in district j of race g and of a certain age, sex, and education, votes for the Republican in a given year. The model applies a shrinkage estimator to impute these probabilities even for demographic combinations for which few respondents exist, respecting the geographic grouping of districts. This first step of our process is similar to the approach in Ghitza and Gelman (2013, 2020).

Our model takes the form

$$\Pr(Y_i = 1) = \text{logit}^{-1}(\alpha_{j[i]} + \beta_{1j[i]}\text{Black}_i + \beta_{2j[i]}\text{Hispanic}_i + \beta_{3j[i]}\text{Other}_i + \boldsymbol{\zeta}_{j[i]}^\top \mathbf{W}_i) \quad (2)$$

where the notation $j[i]$ denotes the district that respondent i resides in; the indicator variables Black_i , Hispanic_i , and Other_i equal 1 if respondent i is of that race (with White as the omitted baseline); and \mathbf{W}_i are individual-level predictors of education, age group, sex, and the interaction of education and race. The function $\text{logit}^{-1}(\cdot)$ is the inverse logit function and transforms an unbounded value to a probability scale between 0 and 1.

Importantly, this model specifies a hierarchical structure to the coefficients. We allow the coefficients to vary across congressional districts and as a function of the demographic composition of the districts. The intercept α is a batch of 435 coefficients with each value representing a district, rather than a single national intercept for all voters. We further encoded geographic hierarchy into these coefficients. Districts are nested within states, and states are nested within divisions, as follows:

$$\alpha_j \sim \mathcal{N}(\alpha_{\text{state}[j]} + \boldsymbol{\rho}^\top \mathbf{V}_j, \sigma_{\text{cd}}^2), \quad \alpha_{\text{state}} \sim \mathcal{N}(\alpha_{\text{division}[\text{state}]}, \sigma_{\text{state}}^2), \quad \alpha_{\text{division}} \sim \mathcal{N}(\alpha, \sigma_{\text{division}}^2)$$

where district-level variables \mathbf{V}_j contains a spline expansion of the Republican Presidential voteshare in district j (Daily Kos 2021) and the proportion of White voters in district j . The coefficients β_1 , β_2 , and β_3 are each batches of coefficients that represent the difference in Republican vote among racial minorities relative to White voters. These also vary by congressional district. Here we specified the coefficient structure so that the values are centered to a national-level coefficient for each race shifted by the contextual effect of the group’s population size in the district.¹ Including the racial composition of the CD incorporates what the ecological inference literature calls linear contextual effects (Online Appendix A.4). The coefficients ζ on other covariates have a similar structure, but with the division-level parameter

¹ For instance the batch of coefficient for Black voters are drawn from a similar geographic hierarchy as the α terms: $\beta_{1j} \sim \mathcal{N}(\beta_{1,\text{state}[j]} + \gamma \mathbf{X}_j, v_{\text{cd}}^2)$, $\beta_{1,\text{state}} \sim \mathcal{N}(\beta_{1,\text{division}[\text{state}]}, v_{\text{state}}^2)$, $\beta_{1,\text{division}} \sim \mathcal{N}(\beta_1, v_{\text{division}}^2)$, such that the coefficients are ultimately centered on a value β_1 , a national-level average difference between Black and White voters. Here \mathbf{X}_j is the proportion of Black voters in district j . The coefficients β_2 and β_3 follow a similar structure.

being centered at 0.

We estimate the model with a Bayesian framework using the `brms` interface (Stan Development Team 2021; Bürkner 2017). We index posterior Monte Carlo samples from the model as $m = 1, \dots, M$ and we retained $M = 2000$ thinned samples across 4 chains. The full specification of the implementation is described in Online Appendix A.

The modeling literature refers to this structure of modeling coefficients as random effects, with α called random intercepts, and β and ζ called random slopes (Gelman and Hill 2006). Random effects use sparse data effectively by shrinking estimates toward a state, regional, or national estimate. If there is no geographic variation, our model reduces to the standard logistic regression with district-level voteshare and national racial group differences (Ghitza and Gelman 2013, 765).

The specification of the model, clearly, can affect estimates. We examined a range of specifications through iterative model building (Gelman et al. 2020). We found that models with race coefficients varying by district lead to more variability in the final estimates than models that let race coefficients vary by region. For example, the choice of specification affected the estimates for Hispanics by 6-10 percentage points.

Survey-assisted Synthetic Target Estimation

We develop a new algorithm for computing the target population. Standard MRP takes the weighted average of the predicted outcome from the hierarchical model, with weights designed to make the sample representative of demographics in the district. Doing so requires knowing the target distribution, or population size, of each [age \times sex \times education] group for each [race \times CD] combination. Unfortunately, no such table of data is available from the United States Census. The Current Population Survey from the Census does not produce statistics at the CD level. The American Community Survey (ACS), the largest survey conducted by the Census, provides a table of the population for each [age \times sex \times education

\times CD] combination and a table of [race \times CD], but not the joint table of the five variables (Online Appendix A.5). In fact, no such weighting target has been provided by public MRP studies.

We therefore develop an algorithm — survey-assisted synthetic target estimation — for fusing together two ACS tables and construct a table of [age \times sex \times education \times race \times CD]. The current recommendation in the MRP literature is either to avoid modeling by assuming independence between the covariates, or to take survey estimates of the joint distribution at face value (Leemann and Wasserfallen 2017; Kestellec et al. 2015). Both the survey data and known marginal constraints inform the joint distribution. Specifically, we estimate a multinomial regression predicting the four categories of education from race, age, and sex on survey data, with a constraint that the implied marginal distribution of education from that regression matches the education composition in each congressional district as reported in the ACS.² The constraint is imposed simultaneously with estimation, instead of applied after fitting the regression (Ghitza and Steitz 2020). This algorithm, available in the `synthjoint` R package, will improve future implementations of MRP and of survey weighting generally as it offers a solution to the limited release of Census tables.

The ACS statistics we use are for the general adult population (i.e., the voting age population), not voters. We therefore repeat the survey-assisted synthetic target estimation and model the turnout rate within each demographic cell. The turnout model involves predicting validated vote among CCES voters and non-voters with education, race, age, and sex, with a constraint that the implied turnout rate matches the actual turnout rate in each congressional district.

Our synthetic target estimation produces a turnout-adjusted [age \times sex \times education] table

² Formally, let Z be a categorical variable for education and \mathbf{X} be the demographic predictors. The marginal distribution of Z and the joint distribution of \mathbf{X} is given by the ACS, but the joint distribution of Z and \mathbf{X} is unknown. A regression of Z on \mathbf{X} in the survey produces fitted probabilities $\hat{\Pr}(Z|\mathbf{X})$ and allows us to compute the joint. We estimate this regression but with the constraint that $\sum_{\mathbf{x}} \hat{\Pr}(Z|\mathbf{X} = \mathbf{x}) \Pr(\mathbf{X} = \mathbf{x})$ matches $\Pr(Z)$ for every district up to a small tolerance. Online Appendix A.6 – A.7 describes the estimation strategy and the specification in this example.

of population sizes for each [race \times CD]. Formally, we index a cell in the table for a given race \times CD combination by $s \in \mathcal{S}_{gj} = [\text{age} \times \text{sex} \times \text{education}]$, so that N_{gjs} denotes the population size of each cell s for racial group g at congressional district j . Each posterior sample m of the estimated coefficients in the hierarchical model provides predicted probabilities of the Republican voteshare in each cell, denoted $\hat{\pi}_{gjs}^{(m)}$. Then, we poststratify cell estimates by weighting them proportional to its associated size in the population:

$$\hat{\tau}_{gj}^{(m)} = \sum_{s \in \mathcal{S}_{gj}} \frac{N_{gjs}}{N_{gj}} \hat{\pi}_{gjs}^{(m)}. \quad (3)$$

This approach does not rely on acquiring a voter file or using estimates of race or education, such as Ghitza and Gelman (2020). Self-reported race is only available in the voterfiles of six states, all of them in the Deep South, and education is never available. Furthermore, the survey-assisted synthetic target algorithm is open-source; it does not rely on proprietary data or algorithms.³ Online Appendix A.3 describes the specifics of this modeling.

Two-way Calibration to Election Results and External Surveys

MRP also assumes ignorable selection into the survey conditional on available covariates (Si 2021). Violation of those assumptions would yield biased estimates (Buttice and Highton 2013). Standard MRP cannot control for a variable that is not measured in the weighting target.

To adjust for selection bias that may remain even after poststratification, we calibrate the MRP estimates from Equation (3) so that the implied district vote matches the actual Republican Presidential voteshare in each congressional district, and the implied vote choice by race matches an external national estimate. For example, we can sum our estimates across racial groups as $\hat{\tau}_j = \sum_{g=1}^G (N_{gj}/N_j) \hat{\tau}_{gj}$, and denote the actual Republican voteshare in district

³ Post-stratification tables for years other than 2016 and 2020 will also be available in the Dataverse (Kuriwaki et al. 2023).

j as τ_j . Here τ_j is a target, and both $\widehat{\tau}_j$ and τ_j are observable. We do not observe a target for our main quantity of interest τ_{gj} , but any observed difference between $\widehat{\tau}_j$ and τ_j can be due to bias in our estimates $\widehat{\tau}_{gj}$ that does not cancel out when aggregated.

Existing work proposes to shift each estimate by a constant so that the implied sum matches an election result (Ghitza and Gelman 2013; Rosenman, McCartan, and Olivella 2023). Unfortunately, such a one-way calibration may make the estimates of the racial gap $\tau_{\text{white},j} - \tau_{\text{black},j}$ even more biased if the estimates of each race-specific voteshare are biased in opposite directions. We find that this is precisely the case in the Florida subset of the CCES we use for validation in the subsequent section.

A two-way calibration adjusts the estimates to a racial group target at the national level (i.e., targets τ_g), in addition to targeting the district level election results. It imposes two sets of additive corrections to the MRP estimates $\widehat{\tau}_{gj}$,

$$\widehat{\tau}_{gj}^{(m)} = \text{logit}^{-1} \left(\text{logit}(\widehat{\tau}_{gj}^{(m)}) + \widehat{\delta}_g^{(m)} + \widehat{\delta}_j^{(m)} \right),$$

where δ_g denotes the racial group correction, and δ_j denotes the geographic correction. The optimal value of the correction minimizes the sum of squared deviances at the national and district level:

$$\begin{aligned} \widehat{\delta}^{(m)} = \arg \min_{\delta} \left\{ \underbrace{\sum_{g=1}^G \frac{N_g}{N} \left(\tau_g - \sum_{j=1}^J \frac{N_{gj}}{N_g} \text{logit}^{-1} \left(\text{logit}(\widehat{\tau}_{gj}^{(m)}) + \delta_g + \delta_j \right) \right)^2}_{\text{Deviation from racial group target for race } g} \right. \\ \left. + \sum_{j=1}^J \frac{N_j}{N} \left(\tau_j - \sum_{g=1}^G \frac{N_{gj}}{N_j} \text{logit}^{-1} \left(\text{logit}(\widehat{\tau}_{gj}^{(m)}) + \delta_g + \delta_j \right) \right)^2 \right\}. \end{aligned}$$

Deviation from geographic target for district j

In other words, for each posterior sample m , we identify the correction factors that make all the weighted sums of adjusted quantities $\widehat{\tau}_{gj}^{(m)}$ match the known election result τ_j and exter-

nal estimate by racial group τ_g as close as possible. We estimate these corrections through numerical optimization.

We use the Presidential voteshare at the congressional district level as district-level targets τ_j , and use the national estimates of vote choice reported by Catalist as the race-level national targets τ_g . The Catalist targets are similar to the National Exit polls. Online Appendix B presents and compares these targets. The Catalist estimates of the Hispanic vote is about 2 to 3 percentage points more Democratic than the CCES before this calibration.

The mean of each posterior sample’s calibrated estimate, $\tilde{\tau}_{gj} = \sum_m \tilde{\tau}_{gj}^{(m)} / M$, represents our best estimate of the vote choice for a specific racial group in a specific district. Our method guarantees that this estimate is (a) computed from a sample that is representative of each congressional district electorate’s joint composition of sex, age, education, and race up to estimation error in the weighting target, and (b) further calibrated so that the estimates for each racial group approximately sum to known election results in the district or at higher levels of geography.

In the 2016 example, the one-way calibration to election results shifts the estimates of White voters’ Trump vote up by 4 percentage points on average, and shifts the estimates among Hispanics up by 1 percentage point, although some districts are shifted in the opposite direction (Online Appendix A.10). The two-way calibration shifts the one-way calibrated White vote up even further towards Trump by 1 percentage point, and pulls the Hispanic vote down by 5 percentage points. These differences suggest that the underlying survey data under-represented Trump voters in most places even after weighting, but that pattern further varied across racial groups.

Evaluation and Validation

It is impossible to validate these methods for vote choice directly, because individual votes are not observed. Voter registration statistics in the state of Florida offer an excellent opportunity

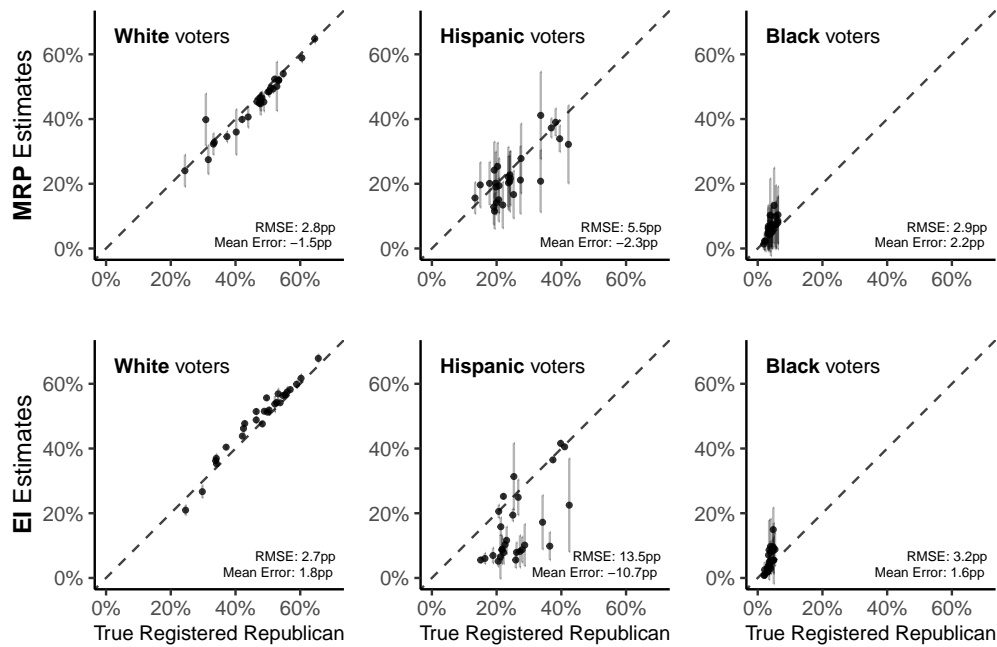
for evaluating the accuracy of our approach (de Benedictis Kessner 2015). Florida and North Carolina are the only two states that record voter registration by party and by race, and their data are publicly available. Party registration in Florida is highly correlated with Trump vote at the precinct level. In the CCES, 83 percent of registered Republicans in Florida report voting for the Republican President, while only 50 percent of registered voters with no party affiliation voted for Trump. We choose to study Florida as the state has substantial Black, Hispanic, and White populations and a large number of districts.

We follow the same methods and post-stratification target as our main analysis and refer to our estimates as the MRP estimates. Instead of modeling the Trump vote in the hierarchical regression stage, we model whether or not the respondent's validated voter registration records them as a Republican registrant. We use the subset of our survey respondents who are in Florida, because unlike the Trump vote, only some states record party registration on their voter files. This means that we partially pool across districts within Florida, which vary in their racial and partisan characteristics.

We also perform a validation with Ecological Inference (EI) using precinct aggregate data, which is the dominant method used to study vote choice by race. We acquired an aggregate dataset recording the racial composition and the party registration composition in each Florida precinct. The average congressional district in this data is represented by 200 precincts, and the median precinct includes 1,300 voters (Details of the dataset are in Online Appendix B.3). We ran a standard multinomial ecological inference model (Collingwood et al. 2016) with the same 4 racial categories as MRP and 3 party categories (Republican, Democrat, and all other parties with Non-Party Affiliated). We run a separate model from each district, using that district's precincts only.

We find that our method produces reasonably accurate estimates and reasonable measures of uncertainty. Moreover, these estimates have better coverage properties than a standard ecological inference method (EI) that has access to data on the entire population aggregated into

Figure 2 – Validation of MRP and Ecological Inference Models in Florida. Each model predicts the proportion of each racial group electorate in each congressional district that are registered Republicans. Statistics in each facet show Root Mean Square Error across districts and the Mean Error (estimate minus truth) across 27 districts. Uncertainty intervals are 80 percent credible intervals. See Online Appendix B.3 for detailed methodology.



precincts. Figure 2 compares Republican registration share for a racial group in a congressional district on the x-axis with the corresponding estimates from our MRP method and traditional EI in the y-axis. Each graph shows the Root Mean Square Error (RMSE), which captures the total error across all districts, and the Mean Error which distinguishes whether these errors systematically tend to overestimate or underestimate.

Our survey-based estimates and EI have similarly-sized error in estimating the Republican registration of White voters, with a RMSE around 3 percentage points. However, the errors go in opposite directions: our survey-based approach tends to underestimate the Republican-ness, while EI tends to overestimate. For Hispanic voters, our estimates outperform EI estimates more than two-fold (an RMSE of 5.5 percentage points vs. 13.5 for EI estimates). Using the survey methods developed here, thus, indicates promise in estimating more accu-

rately the voting preferences for groups. Using ecological inference can lead to substantively different conclusions about racial cohesion and polarization, and tends to show higher levels of polarization than does the survey data.

Survey estimates using hierarchical regression quantify uncertainty reasonably well, while the confidence intervals for EI are too tight. For all four racial groups, the 80 percent confidence intervals of each of our survey-based estimates cover the true value in 78.7 percent of 108 district \times race combinations. The 80 percent confidence intervals of the EI estimates cover the true value in only 38 percent of the 108 estimates. The lack of coverage for EI owes partly to the inaccuracy of EI for Hispanic registrants and partly to the fact that the standard errors are overly tight for Black and White registrants. Appropriately calibrating survey data using our three-step approach, then, can lead to more accurate estimates and inferences than traditional survey data with weights or aggregate data methods such as EI. We leave to further research the questions of diagnosing issues with EI and how applications of that method might be improved. Although our method, and MRP in general, does have its own estimation errors, we rely on individual-level data and can provide reasonably accurate estimates of the relationship between race and vote choice.

Our validation using data from Florida is the best setting for a comparison of the accuracy of EI and our survey-based approach. Most other states lack voter file data on race and party. However, in Online Appendix B.4 we obtain EI estimates for all 435 districts based on election data from the Voting and Election Science Team (2020) and precinct-level racial group proportions for the 2020 Census prepared by McCartan et al. (2022), and show where our estimates differ. In the average district, EI differs from our survey estimates by over 10 percentage points for White voters and Black voters, and over 20 percentage points for Hispanic voters.

THE POLITICAL GEOGRAPHY OF VOTE CHOICE BY RACE

The estimates offer a nuanced picture of the geography of racial voting preferences. Consistent with the group voting literature, Black voters overwhelmingly prefer Democrats and that preference does not appear to vary by geography. Consistent with the perspective that geography and context matter, the preferences of White and Hispanic voters vary considerably with place.

The Racial Gap at the District Level

The gap between White and non-White voting preferences varies considerably across the United States. Figure 3 maps the racial gap between White and all non-White voters (defined in Equation (1)) from our final estimates. Because congressional districts are roughly equal in population but vary in land area, we use the cartogram by Daily Kos that sizes districts equally while approximating each district's location within a state.⁴

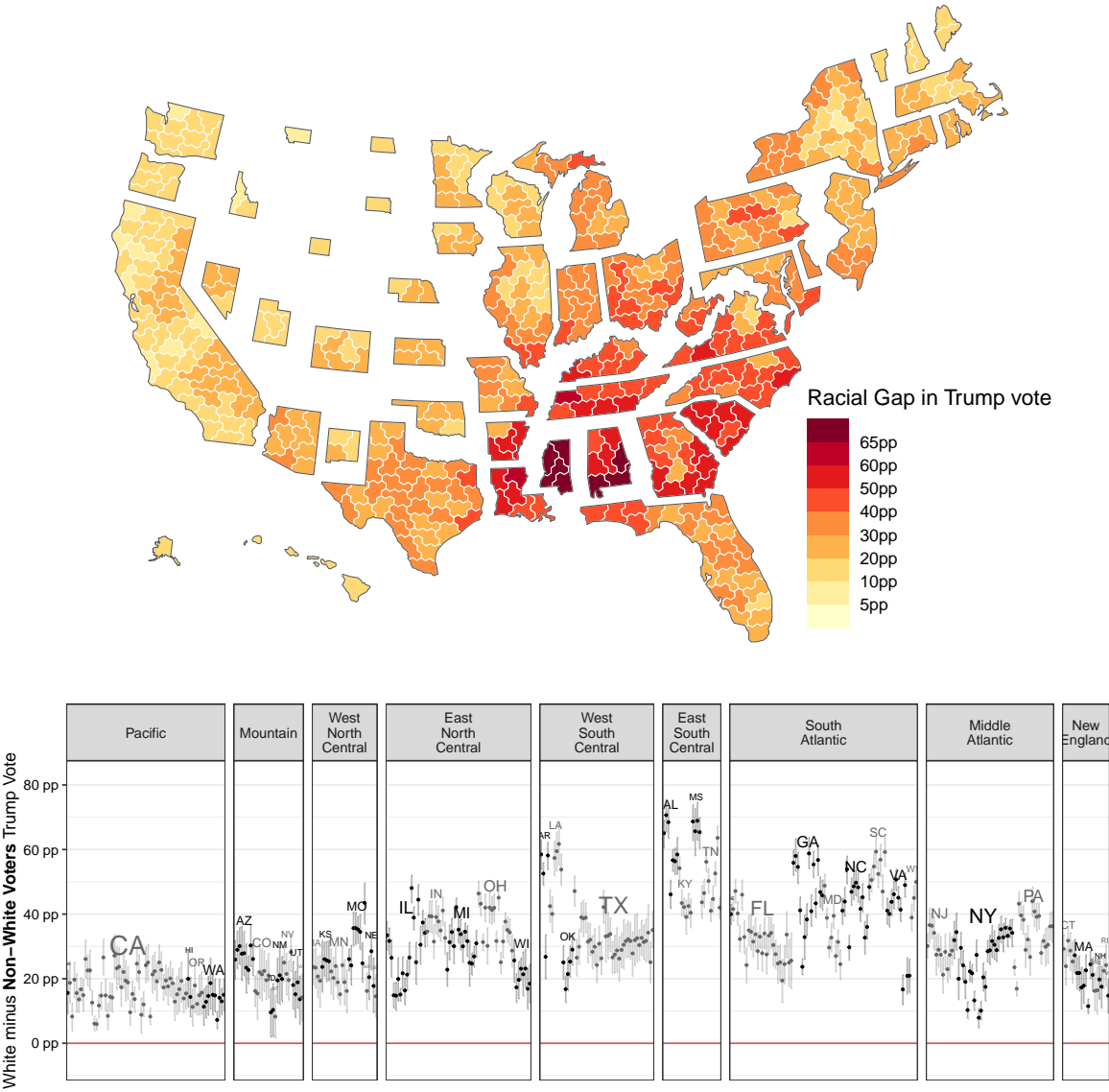
The dark red area of the map, centered in Louisiana, Mississippi, Alabama, are the CDs that exhibit the highest racial gaps, in excess of 60 percentage points. The seven highest racial gap estimates all appear in Mississippi and Alabama. In these districts, the gap between White and non-White voters is close to 70 percentage points.

The districts with the lowest racial gap tend to be urban areas. In 2016, the five least polarized districts include NY-12 (East side of Manhattan and parts of Queens⁵), CA-12 (San Francisco, represented by Speaker Pelosi), CA-13 (Oakland, Rep. Barbara Lee (D)), and WA-07 (city of Seattle, Rep. Pramila Jayapal). In these districts, the racial gap ranges from 6 to 8

⁴ As a result, locations are not exact. For example, districts in New York City are placed closer to upstate New York because there are too many NYC districts to fit in the southeast corner. The correspondence between polygons in this map and each district is given in Online Appendix C. Substantively interpretable names of those districts (e.g. East Bronx) are given in the estimates in Dataverse, <https://doi.org/10.7910/DVN/MAZNJ6>.

⁵ Throughout the paper, we refer to the geographies and representatives as of 2021.

Figure 3 – Racial Gap between White and Non-White Voters. Colors in each congressional district show the point estimate of the estimated racial gap in the 2016 Republican vote between Whites and Non-Whites in that district (Equation (1)). The map is displayed in the Daily Kos cartogram (<http://dke1.ec/map>): Districts are placed in their approximate place within the state, and states are sized by their population. In the bottom panel, each district point estimate is represented by a point on the vertical axis. Districts are sorted first by Division in facets (from West to East), then by state (alphabetical within Division), then by district number. States are annotated and alternatively colored black and gray for visual clarity.



percentage points. The distribution of the racial gap is skewed, with large racial gaps in the Deep South pulling the average district's racial gap to be about 2 percentage points higher than the median district (Online Appendix C).

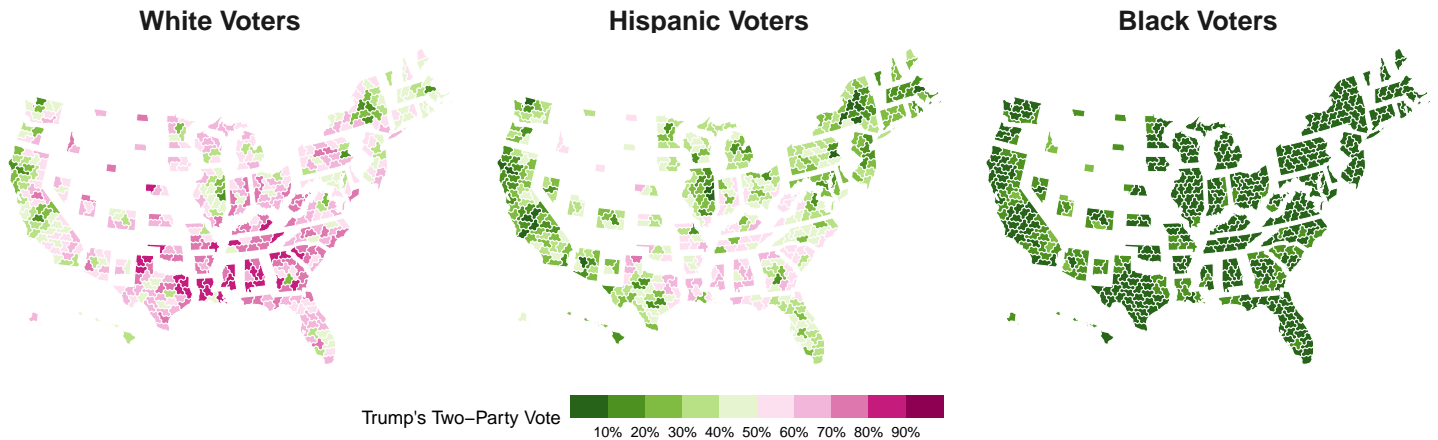
Although the deep red parts of the map are visually striking, the interpretation of the gap and its magnitude depends on the context. A racial gap of 35 percentage points would arise from an electorate in which 70 percent of White voters vote for the Republican and 35 percent of non-White voters vote for the Republican candidate. Such a situation exhibits racially polarized voting and may require creation of a majority minority district under Section 2 of the VRA. However, a CD in which 95 percent of non-Whites vote Democratic and 60 percent of Whites vote Democratic has a similarly large racial gap (35 points) but is not racially polarized, because a majority of both groups prefer the same candidate.

Disaggregation by district highlights the variation within a state. For example, Chicago and the northern parts of Illinois have the lowest racial gap estimates in the Midwest but districts in Southern Illinois adjacent to St. Louis have racial gaps as high as those in Tennessee and TX-01 (northeastern Texas, represented by Rep. Louie Gohmert). A similar pattern occurs in Virginia, where the DC suburbs of VA-08, VA-10, and VA-11 have low racial gap values similar to the Boston suburbs, but other Virginia districts have racial gap values that are in the top quarter of all districts.

From coast to coast, the racial gap rises in the Midwest and the South, and falls again in New England. The bottom panel in Figure 3 shows the same estimates of differences on a common vertical axis, and shows the 80 percent credible interval for the uncertainty in the estimates. Almost all districts have intervals that do not include zero.⁶

⁶ We compute the 80 percent credible interval of the difference in $\tilde{\tau}_{\text{white},j}$ and non-White vote $\tilde{\tau}_{g,j}$ by taking the 10th and 90th quantiles of the 2000 posterior sample estimates of differences. The posterior distribution also allows us to account for the dependency between the White vote and the minority vote conditional on a district, which could be positively or negatively correlated and violate the standard assumption in two-sample t-tests.

Figure 4 – 2016 Republican Vote by Racial Group in Congressional Districts. Each district is colored by the estimated two-party voteshare for the Republican Presidential candidate.



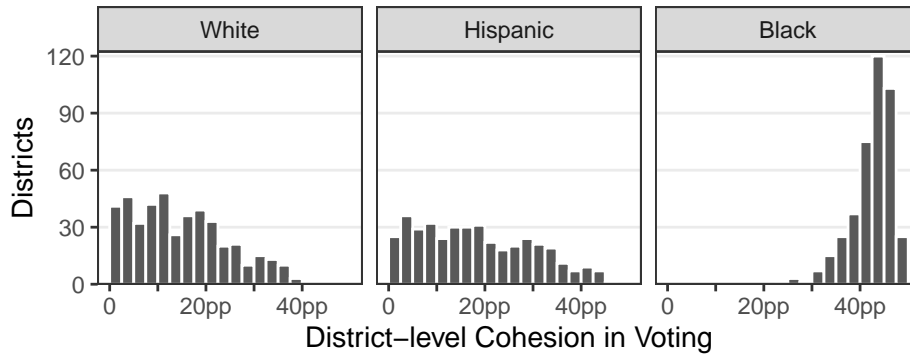
Group Cohesion

The variation in the racial gap between White and non-White voters is largely driven by White voters and Hispanic voters. Black voters cohesively vote for the Democratic party in all congressional districts. Here we examine the cohesiveness of each group nationally and across districts, states, and regions. Figure 4 decomposes the cohesion estimates of vote choice into the two-party voteshare by the three major racial groups.

The voting tendencies of White voters in the urban areas of California, New York, and Illinois are a stark contrast to White voters in the Deep South, where over 80 percent of White voters voted for Trump. Hispanic voters also exhibit high levels of variation. Hispanic voters in Southern Florida are more Republican than the Hispanics in San Francisco or Chicago. Among districts where over 40 percent of the electorate is Hispanic, the district in which the Hispanic voters are the most Republican is FL-25, which has the largest Cuban American population in the US and is represented by Rep. Díaz-Balart (R).

We use the term *cohesion* to refer to the voting behavior of a group (Pildes 2001; Atsusaka

Figure 5 – District Level Cohesion. The range of district-level cohesion of each of the three racial groups, quantified as the absolute distance from 50 percent of the point estimate of the group’s two-party support in the district. States where the Black or Hispanic population is less than 2 percent of the electorate are not analyzed.



2021). Cohesion equals the absolute deviation of the point estimate of the percent of a group that votes for the Republican from 50 percent. For example, White voters voting 85 percent for Donald Trump will have the same cohesion value as Hispanic voters voting 15 percent for Trump: 35 percentage points. Figure 5 shows the range of these district cohesion scores for White, Hispanic, and Black voters. Black voters are highly cohesive in their preferred candidate, while Hispanic and White voters show considerable variation in the degree of cohesion across CDs. In 385 districts, Black voters’ cohesion is over 0.35 (i.e., 85 percent vote for one party); White voters have the same level of cohesion in only 13 districts and Hispanic voters in only 35. Using ecological inference on precinct-level election data produced higher cohesion estimates for White voters (by 6 percentage points on average) and Hispanic voters (12 percentage points), with lower cross-district variance.

Our goal is to measure the magnitude and patterns of racial group preferences, not to explain their source.⁷ Nonetheless, it is worth noting that there is a clear urban-rural gradient evident in Figure 4. Our estimates reveal a substantial difference between the most urban and most rural CDs among White voters and among Hispanic voters. White voters in urban

⁷ Work that seeks to explain geographic or racial variations in preferences highlight place-based identities (Cramer 2016; Munis 2020), racialized social constraint (White and Laird 2020), and political incorporation (Masuoka et al. 2018; Alvarez and Bedolla 2003).

districts are more Democratic than White voters in suburban districts, and even more so than White voters in rural districts. That pattern is consistent with research on group preferences as a function of ZIP code density and distance to large cities (Gimpel et al. 2020), and it is consistent with the notion that voting rights law needs to be narrowly tailored to the voting patterns in particular areas. Even still, within urban, suburban and rural CDs, a 20-percentage point difference between White and Hispanic voters remain. Substantial racial group differences, then, are not just a matter of where people live.

Districts are, in turn, nested within states and regions. Table 1 shows our point estimates of vote by race at the national level, then separates this estimate to the four U.S. Census regions, then by the nine Census divisions, and then by each of the 50 states. Our statewide and national estimates differ from other surveys such as the National Exit Polls by only a few percentage points (Online Appendix B).⁸ Our procedure increases the precision, i.e. decreases the standard error, in subgroup estimates, relative to the raw or post-stratification weighted survey data at the CD or state level. The standard error of our state-level estimates ranges from about 0.005 to 0.10, inversely proportional to the size of the group (Online Appendix C.4). For example, there are roughly $n = 300$ Alabama voters in our survey data, a quarter of whom are Black. That implies a standard error of about 0.033 for White voters in Alabama and 0.06 for Black voters when the sample is taken as-is with no modeling. The comparable standard error of our modeled estimates is about 0.01 for White voters and 0.023 for Black voters.

In 2016, the White - Non-White gap was 37 percentage points: 59 percent of White voters voted Trump, whereas only 29 percent of Hispanic voters and 7 percent of Black voters voted for Trump. But Whites in the Northeast were 8 points less likely to support Trump than Whites in the Midwest, and 18 points less likely than Whites in the South (Table 1(a)). Within the Northeast, moreover, voting among Whites differed by 10 points between New

⁸ An exception is a 10-15 point difference with Latino Decisions and the Collaborative Multiracial Post-Election Survey (CMPS) on the Republican vote of Hispanics, which we discuss in Appendix B.

England (45 percent) and the Middle Atlantic states (54 percent) of New York, New Jersey, and Pennsylvania (Table 1(b)). At the state level we see further variation. The Republican voting patterns among White voters in the Deep South states are different from those the peripheral South states (McKee and Springer 2015), as well as different within other regions.

Overall, Table 1 shows how measuring the racial gap only at the state or national level masks important variation in large states. With state-level estimates, White voters appear solidly Republican: In all states but five in New England and four in the Pacific West, the majority of White voters vote for Donald Trump in 2016. But state groupings mask other differences in congressional district constituencies like the urban-rural divide.

An analysis of variance (ANOVA) is an ideal framework for summarizing the relative importance of racial group differences and nested geographic group differences (Gelman and Hill 2006, ch.22). An ANOVA can partition the overall variation in our estimates by the variation explained by a race-level average, a geography-level average, or the interaction of race and geography. Recall that we have $G \times J \times M$ estimates of vote choice $\tilde{\tau}_{gj}$ for each racial group g and congressional district j . We first consider the following three-part decomposition,

$$\tilde{\tau}_{gj} = \mu + \phi_g + \eta_j + \gamma_{gj} + \varepsilon_{gj}, \quad (4)$$

where we partition the estimates of vote choice into the racial group component (ϕ), the geography component (η), and their interaction (γ). The error term ε represents sampling variation. By decomposing district-race vote shares in this way, we are defining the geography component as a single Republican voteshare for a given district j that does not vary by racial group. It can be thought of as the normal partisan lean of the entire district. The definition is agnostic as to whether this partisan lean is due to factors such as rurality, or place-based identities.

We are interested in how much variation in the patterns we found is explained by each

Table 1 – 2016 Republican Vote by Racial Group in Regions, Divisions, and States

(a) By Region

	Non-Whites				All
	White	All	Black	Hisp.	
Northeast	51	17	4	20	43
South	69	23	8	37	55
North Central	59	19	5	29	53
West	50	26	9	25	42
National	59	22	7	29	49

(b) By Division, nested within Region

	Non-Whites				All
	White	All	Black	Hisp.	
Northeast					
New England	45	19	3	19	41
Middle Atlantic	54	16	4	21	44
South					
South Atlantic	64	21	7	37	51
East South Central	77	16	7	50	64
West South Central	72	30	9	37	58
North Central					
East North Central	57	17	5	25	51
West North Central	61	28	6	39	57
West					
Mountain	59	33	13	29	53
Pacific	44	24	9	23	36
National	59	22	7	29	49

Note: Numbers show estimated two-party voteshare for the Republican Presidential candidate for each geography and racial group rounded to the nearest full percentage point. Columns represent, from left to right, the Republican vote among White voters, all non-White voters, Black voters, (any-part) Hispanic voters, and all racial groups in the geography. Numbers in parentheses are those where we estimate the racial group comprises only 2 percent or less of the geography's electorate. See Online Appendix C.4 for standard errors of the state-level point estimates.

(c) By State, nested within Division

	Non-Whites				All
	White	All	Black	Hisp.	
New England					
Connecticut	50	20	4	24	43
Maine	49	33	(4)	(26)	49
Massachusetts	40	16	2	15	36
New Hampshire	51	32	(4)	(25)	50
Rhode Island	45	21	3	21	42
Vermont	36	21	(2)	(15)	35
Middle Atlantic					
New Jersey	55	20	4	22	43
New York	50	15	4	18	39
Pennsylvania	58	15	3	28	51
South Atlantic					
Delaware	54	14	5	26	44
Florida	63	27	6	37	51
Georgia	73	19	8	44	53
Maryland	51	13	6	25	36
North Carolina	66	19	8	39	52
South Carolina	73	17	10	50	58
Virginia	60	19	7	32	47
West Virginia	74	29	9	(54)	72
East South Central					
Alabama	83	17	7	60	65
Kentucky	71	21	3	40	66
Mississippi	84	15	10	(60)	60
Tennessee	74	15	3	44	64
West South Central					
Arkansas	74	23	8	45	65
Louisiana	81	22	13	54	61
Oklahoma	73	48	6	48	70
Texas	70	31	6	36	55
East North Central					
Illinois	51	15	5	19	41
Indiana	65	23	7	38	60
Michigan	57	17	5	31	51
Ohio	61	16	6	(34)	55
Wisconsin	54	23	3	31	51
West North Central					
Iowa	57	35	7	43	55
Kansas	65	37	8	43	61
Minnesota	53	24	4	27	50
Missouri	66	20	6	(43)	60
Nebraska	66	39	8	46	64
North Dakota	71	55	(12)	(52)	70
South Dakota	67	53	(10)	(47)	66
Mountain					
Arizona	60	28	13	23	52
Colorado	52	31	12	30	48
Idaho	69	59	(24)	53	69
Montana	62	54	(16)	(38)	62
Nevada	57	32	12	28	49
New Mexico	55	34	(14)	32	46
Utah	65	48	(18)	39	63
Wyoming	77	63	(27)	59	76
Pacific					
Alaska	62	47	16	45	59
California	43	23	9	22	34
Hawaii	42	26	9	23	33
Oregon	46	29	(6)	27	44
Washington	45	28	7	30	42
National	59	22	7	29	49

component. In ANOVA these relative importance measures are given by

$$\widehat{\kappa}_{\text{race}} = \frac{J \sum_{g=1}^G \widehat{\phi}_g^2}{\text{TSS}(\widehat{\tau})}, \quad \text{and} \quad \widehat{\kappa}_{\text{cd}} = \frac{G \sum_{j=1}^J \widehat{\eta}_j^2}{\text{TSS}(\widehat{\tau})},$$

where TSS is the total sums of squares $\text{TSS}(\widehat{\tau}) = \sum_g \sum_j (\widehat{\tau}_{gj} - \widehat{\mu})^2$. The components are estimated via OLS with the sum-to-zero constraint on each group of coefficients, $\sum_g \phi_g = \sum_j \eta_j = \sum_j \sum_g \gamma_{gj} = 0$.⁹ We interpret the κ term for each component as a proportion because they sum to 1.¹⁰ A large value of $\widehat{\kappa}_{\text{race}}$ would indicate that the variation in the final estimate is largely explained by a single national racial group difference in Republican vote-share that does not vary with geography, while a large value of $\widehat{\kappa}_{\text{region}} + \widehat{\kappa}_{\text{state}} + \widehat{\kappa}_{\text{cd}}$ would imply that a region, state, or district-level voteshare explains more of the total variation in the estimates. A large value of $\widehat{\kappa}_{\text{residual}}$ would imply that much of variability is posterior estimation uncertainty rather than anything systematic.

The simple model in Equation (4) can be made more elaborate to partition the geographic component η to states and regions, in addition to districts. This amounts to estimating separate terms for state and region components, e.g. $\eta_{\text{state}[j]}$ and $\eta_{\text{region}[j]}$. If a single value for an entire state or entire region is sufficient to entirely explain district-level geographic variation, all of the variation previously attributed to η_{cd} will shift to the variation explained by η_{state} .

The ANOVA results are summarized in Table 2. Model (1), corresponding to Equation (4), shows that about 60 percent of the total variation in the district- and race-level vote is explained by a national race pattern, 28 percent is explained by geography, and the interaction of the two (i.e., differences between races that vary depending on the geography) explains the

⁹ For computational tractability we use every tenth sample from M total.

¹⁰ We define the quantity for the interaction term $\widehat{\kappa}_{\text{race} \times \text{cd}}$ and for the residual similarly,

$$\widehat{\kappa}_{\text{race} \times \text{cd}} = \frac{\sum_{g=1}^G \sum_{j=1}^J \widehat{\gamma}_{gj}^2}{\text{TSS}(\widehat{\tau})}, \quad \text{and} \quad \widehat{\kappa}_{\text{residual}} = \frac{\sum_{g=1}^G \sum_{j=1}^J \widehat{\varepsilon}_{gj}^2}{\text{TSS}(\widehat{\tau})},$$

where $\widehat{\varepsilon}_{gj} = \widehat{\tau}_{gj} - \widehat{\mu} - \widehat{\phi}_g - \widehat{\eta}_j - \widehat{\gamma}_{gj}$. Then the four quantities sum to one, $\widehat{\kappa}_{\text{race}} + \widehat{\kappa}_{\text{cd}} + \widehat{\kappa}_{\text{race} \times \text{cd}} + \widehat{\kappa}_{\text{residual}} = 1$.

Table 2 – Proportion of Variance Explained by Race and Geography. Each column represents an ANOVA model estimated from a tenth of the posterior sample estimates of the district- and race-level vote. Proportions show the fraction of variance in the outcome explained by each type of variable ($\hat{\kappa}$), and sum to 1. The last two rows presents the variance across samples (or the total sums of squares divided by the sample size) and the number of posterior samples used.

		Fraction of Variation Explained				
		Across Races		Within Race		
		(1)	(2)	White (3)	Black (4)	Hispanic (5)
Race / Ethnicity						
	Racial Group	.60	.60			
Geography						
	Congressional District	.28	.15	.54	.30	.39
	State		.06	.19	.14	.16
	Region		.07	.26	.18	.23
Race \times Geography						
	Race \times Congressional District	.07	.03			
	Race \times State		.02			
	Race \times Region		.03			
Sampling Variation						
	Residuals	.05	.05	.01	.38	.22
	Total Variance	0.056	0.056	0.027	0.0034	0.029
	Samples	348,000	348,000	87,000	87,000	87,000

remaining 7 percent. Model (2) decomposes the geographic component into districts (nested within states), states (nested within regions), and regions. Congressional districts explain 15 percent of the variation above and beyond larger geographies, whereas states explain only 6 percent and regions explain 7 percent. Estimation uncertainty accounts for about 5 percent of the variation in our data.

The weight of geography (i.e., CD, state, and region) also varies within each racial group. Models (3) - (5) use the estimates of one racial group at a time so that there is no race-level variation. The penultimate row shows that the total variation in the Trump vote among White and Hispanic voters is more than five times larger than the total variation observed for Black

voters. In all racial groups, the district level accounts for more than twice as much variability in the total estimates than state or region-level averages. White and Hispanic voters vary more than Black voters, and the bulk of the variation occurs at the district level.

At a high-level, then, differences in the average vote choice of racial groups nationwide explains twice as much of the systematic variation in voting as does the CD, state, or region. Sixty percent of the variation in the vote of racial groups within a particular congressional district can be accounted for by the average national vote of the groups. The normal vote of a CD, state, or region explains 30 percent of the remaining variation. A final 10 percent of the variation is due to the fact that voting patterns of a given racial group change by geography. In other words, while we do find substantial sub-state variation, geography (i.e., CD, state, and region) explains at most a third of the total variation across all district-race combinations.

Ecological inference estimates yield different conclusions on the explanatory power of geography. We applied the same ANOVA modeling to the posterior sample of ecological inference estimates, which were generated from applying EI one district at a time, described in Online Appendix B. The national race component explains roughly the same amount of variation across both methods (both .52 in 2020). However, state, region, and district explain twice as much of the variation using MRP than using EI (.34 versus .15). Because MRP models district level variation as a random effect, it allows for more efficient estimation of the district component κ_{cd} . EI implemented at the district level forces that variation onto the interaction component $\kappa_{\text{race} \times \text{cd}}$. One may improve EI by allowing partial pooling across districts, but that is not conventionally done.

Our analysis examines a group's voting behaviors, rather than beliefs, ideologies or issue preferences. Our findings reaffirm the lack of variation in party vote among Black voters, but there are certainly diverse dynamics within that group beyond party choice (White and Laird 2020; Jefferson 2023). Even still, we examine correlates of ideology or belief, and the methods developed here open up opportunities to model any survey item to explore how

ideology or public opinion may vary within groups.

Groups and Coalitions

Finally, we turn to the potential of groups to form coalitions. The emerging voting power of Black and Hispanic groups over the past three decades and the cohesiveness of these groups raises the possibility that the two groups may vote sufficiently strongly together in general elections to be able jointly to elect their preferred candidates (Axelrod 1972; Grofman, Handley, and Lublin 2000; Barreto, Collingwood, and Manzano 2010; Atsusaka 2021). The demographic realities of contemporary America are putting pressure on voting rights law to accommodate the possibility of multi-racial coalitions (Pildes 2001).

Coalition voting can be operationalized two ways. First, how often do the majority of each group of minority voters prefer the same candidates? The point estimates underlying Figure 4 indicate that in 356 out of 435 CDs (82 percent), a majority of Black voters and a majority of Hispanic voters voted for the same candidate. Majorities of White and Hispanic voters voted for the same candidate in 216 CDs (49 percent), and majorities of White and Black voters voted for the same candidate in 136 (31 percent).

Second, which group or groups are pivotal (Ingham 2019; Snyder 1989)? We identify the pivotal racial group or groups in each CD using the 2016 Presidential vote. We divide each congressional district into whether it reports more votes won by the Republican, Donald Trump, or the Democrat, Hilary Clinton. We then use the point estimates of the vote shares of the Democratic and Republican candidates for each racial group in each CD, and then calculate the total vote for each group that went to the Democrat and to the Republican.¹¹ Using those totals we determine which groups were pivotal in giving the winning candidate her or his majority. We distinguish these districts in Figure 6 and tabulate the frequency of

¹¹ Our hierarchical regression model estimates a voteshare. We multiply this with the total number of voters for the race-district combination, estimated in our target distribution.

Table 3 – Typology of Racial Coalitions. Ability of party supporters in each racial group to win a majority of votes either on its own or as a coalition. Republican vs. Democratic CDs refers to the party that won the two-party Presidential vote in each congressional district.

Racial Group Composition for Sufficient for a Majority	Democratic CDs	Republican CDs
White voters alone	20	179
Black voters alone	10	0
Hispanic voters alone	2	0
Black and Hispanic voters	18	0
Black, Hispanic, and Other non-White voters	15	0
White and non-White	140	51
Total	205	230

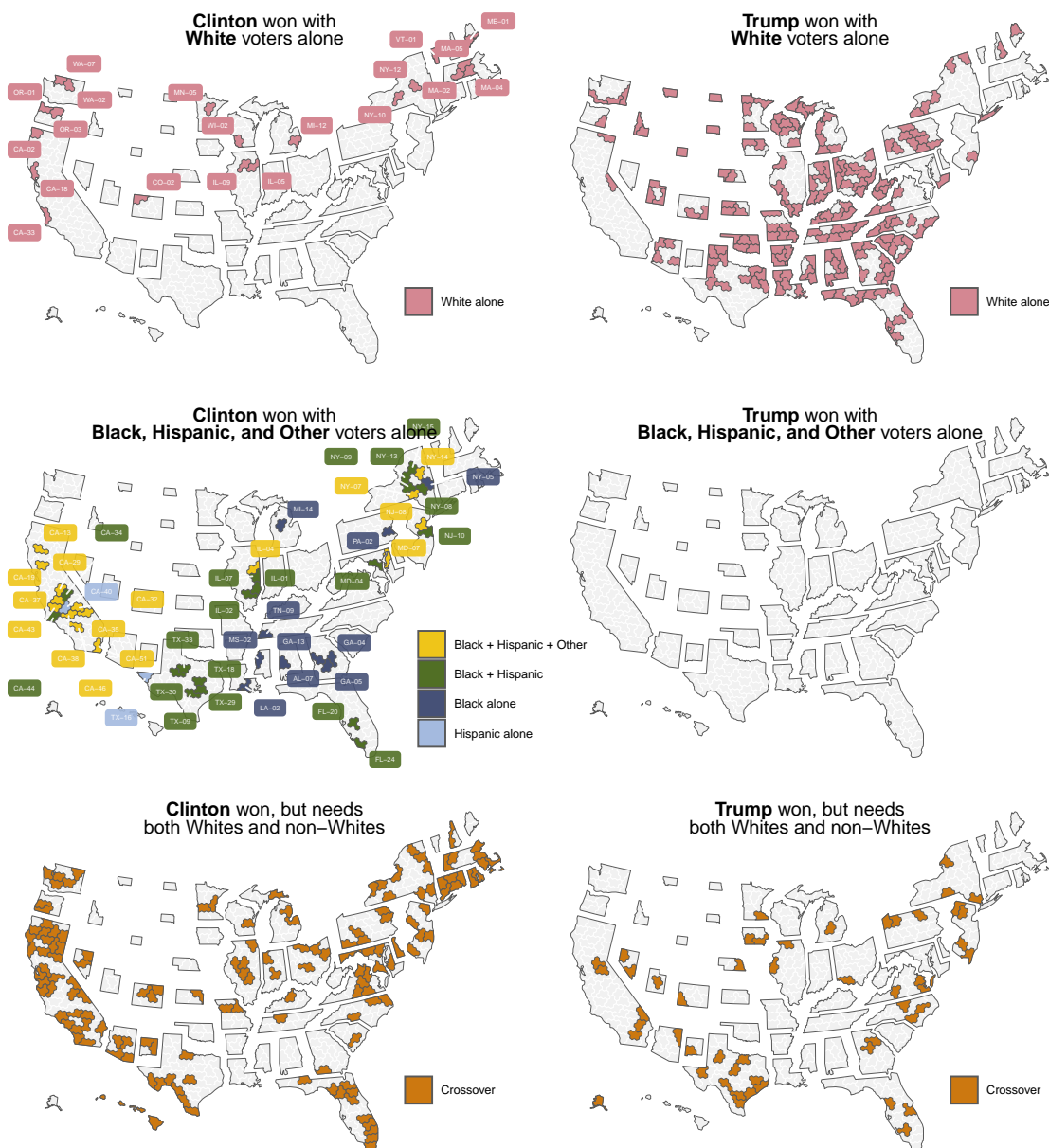
each type in Table 3.

The striking finding in Table 3 is that minority voters are pivotal in more than half of the districts. In 236 of the 435 CDs, minority voters, either one group singly or together with White voters, were pivotal in determining which candidate won the majority of the vote. For example, in UT-04 which includes Salt Lake County, White voters comprise close to 80 percent of the electorate but voted sufficiently for Clinton that White Trump voters alone fall short of a majority. Given their large population, White voters are often part of a pivotal set of racial groups as well (Cohn 2016): their vote is needed in 390 districts for the winning candidate.

Among the 236 districts where minority voters are part of a pivotal coalition, 45 are those in which minority voters are pivotal without White voters. Black voters are large and cohesive enough to deliver a majority of the district vote “alone” (i.e., without relying on other racial groups) in 10 districts, and Hispanics voters are pivotal alone in 2 (Table 3). In a separate 33 districts, minority groups (including Asian Americans, Native Americans, and multi-racial voters), voted with sufficient cohesion to account for a majority of the votes. The center-left panel of Figure 6 identifies these districts. All 45 of districts were won by Hillary Clinton, the Democrat. In no district did Trump’s majority rest solely on non-White voters.

In 199 of the 435 districts, White voters alone are large enough to deliver the majority

Figure 6 – Coalitional and Crossover Districts. 435 congressional districts in the 2016 Presidential election are classified into one of 6 types listed in Table 3. Maps on the left show districts won by the Democrat, and maps on the right show districts won by the Republican. Districts are placed in the same map as Figure 3.



of the votes cast for the winning candidate. The overwhelming majority (179 out of 199) of these heavily White districts voted for Trump.

These statistics suggest that both parties rely on coalitions of White and minority voters. In a fifth of districts Trump won in 2016, Trump needed support of both White and non-White voters to win (Table 3). The Democratic party relies much more on the votes from multiple racial groups. Among the 205 districts won by Clinton in 2016, her majorities relied on a coalition of White and minority voters in 140 districts. Clinton won only 20 districts in which the White vote alone was sufficient to win a majority. This is the multi-racial electoral context that Pildes (2001) foresaw, and it presents new complexities for the crafting and application of voting rights laws. However, the degree of reliance is asymmetric. White voters alone were sufficient to win a majority of votes in districts Trump won, while in the majority of districts Clinton won, Democrats needed votes from both White voters and minority voters (Table 3).

CONCLUSION

This study offers the first set of estimates of racially polarized voting at the congressional district level for all 435 congressional districts and 50 states. Our findings immediately inform debates over persistence of importance of race and regionalism in US elections, specifically the extent to which the racial divide in the US is a national phenomenon or is regionally concentrated. We find national division in the average vote across racial groups that explains 60 percent of the variation in our estimates. Nationally, 22 percent of non-White voters voted Republican while 59 percent of Whites voted Republican in 2016 – a 37 point difference. There are, however, significant cross-state and within-state variations in group voting behavior as well. Black voting behavior is far more consistent across districts and states than Hispanics or Whites. Hispanic and White voting behavior varies considerably. For instance,

only 40 percent of White voters voted for the Republican in Massachusetts, compared to 86 percent in Mississippi. The structure of the variation we find within groups across the CD, state, and regional levels is analogous to that of Erikson, Wright, and McIver (1989), which finds ideological variation within the Republican and Democratic party across states.

Our findings show the complex dynamics of racial group politics in the United States. Differences in election outcomes are explained by both differences across groups and the aggregate differences by geography. And through analyzing patterns of vote share, within-district polarization, and group size, we find that most congressional districts were won either with a multi-racial coalition or a significant cross-over vote of either whites or minorities in 2016. These group dynamics are present even though White and non-White voters are polarized nationally. This fact points to an increasingly important reality for both major political parties in the United States: support from minority voters is a necessity for both parties to win a congressional majority. Demographic trends will only increase the importance of Asian, Black, and Hispanic votes within both parties.

As the voting behavior of the United States electorate shifts, so too must the laws and policies that prevent racial vote dilution. Our analysis supports both the need for federal voting rights law and the need for the narrow tailoring of federal and state rules to specific areas in line with the principles of federalism. Many of the divisions across racial groups that existed in the past remain. Those divisions do, however, vary across regions and states and, even, within states. Majority minority districts may be required, then, in some areas within a state but not others. Our study offers a new assessment, informed by individual-level data, of where such districts may be required to protect minority voters.

Our article also offers innovations in the analysis of survey data that open the possibility of using individual data where only aggregates were available before. The modifications to the existing MRP procedure in our framework can integrate more survey data and aggregate statistics together, generating more reliable estimates of group voting behavior at different

levels of geographic aggregation. Use cases for these methods abound in the social sciences. We have focused on racial voting patterns at the level of districts and states, but this approach could easily be applied to other geographies, such as metropolitan areas, counties, and cities, or to other demographic groups. For example, the tools developed here allow researchers to distinguish specific cultures or nations of origin of more precisely defined groups, such as Mexican Americans, Puerto Ricans, and Cuban Americans, who are often combined under the label Hispanic or Latinx.

There is much potential to extend these methods to more difficult settings with even sparser data, but we also think that such applications must be interpreted with care, must be attuned to the particular problem at hand, and must be validated whenever possible. Important extensions to the method include accounting for geographic adjacency of districts (Morris et al. 2019), and the integration of multiple surveys to increase the sample sizes of particular groups (Frasure-Yokley et al. 2020; Barreto et al. 2018). Integrating other surveys will require extending the methods presented here to reflect the surveys' timing, mode, question-wording, and sampling that could confound the interpretations of differences. There is also a practical implication for ecological inference methods. Pooling across all districts as in a nationwide EI may violate constancy assumptions, but running separate EI methods by district can underestimate the shared explanatory power of cross-district geography. Implementing partial pooling in EI methods may improve estimation.

Finally, our analysis speaks to one of the most important emerging problems in voting rights law. The Voting Rights Act (VRA) of 1965 was developed in an era when politics was often Black versus White. Today, Hispanics are the second largest racial group and Asians are growing quickly. Should the Voting Rights Act require minority coalition districts, and if so, where (Pildes 2001)? Our results show that in most districts the majorities of Black and Hispanic voters both support the same party. That does not mean that minority coalition districts can always be drawn, but there is considerable potential for such districts throughout

the United States. This approach to representation would more accurately reflect the realities of racial voting patterns in American politics today.

DATA AVAILABILITY STATEMENT

Documentation, data, and code that reproduce the findings of this study are openly available in the *American Political Science Review* Dataverse at <https://doi.org/10.7910/DVN/VX5N1V>. Similar estimates for years outside the scope of this study are available via Dataverse <https://doi.org/10.7910/DVN/MAZNJ6>. Software to implement our methods are publicly available through the R packages `ccesMRPprep`, `ccesMRPrun`, and `synthjoint`. The packages are also archived in the Dataverse replication repository.

FUNDING STATEMENT

This research was supported by NSF grant 1926424.

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

HUMAN SUBJECTS

The authors affirm this research did not involve human subjects.

REFERENCES

- Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018. *Deep Roots: How Slavery Still Shapes Southern Politics*. Princeton, NJ: Princeton University Press.
- Alvarez, R. Michael, and Lisa Garcia Bedolla. 2003. "The foundations of Latino voter partisanship: Evidence from the 2000 election." *The Journal of Politics* 65 (1): 31–49.
- Ansolabehere, Stephen, Samantha Luks, and Brian F. Schaffner. 2015. "The perils of cherry picking low frequency events in large sample surveys." *Electoral Studies* 40:409–410.
- Ansolabehere, Stephen, Nathaniel Persily, and Charles Stewart. 2009. "Race, region, and vote choice in the 2008 Election: Implications for the future of the Voting Rights Act." *Harvard Law Review* 123:1385.
- . 2012. "Regional differences in racial polarization in the 2012 Presidential Election: Implications for the constitutionality of Section 5 of the Voting Rights Act." *Harvard Law Review* 126:205.
- Ansolabehere, Stephen, and Douglas Rivers. 1995. "Bias in Ecological Regression." *MIT Department of Political Science Manuscript*.
- Atsusaka, Yuki. 2021. "A Logical Model for Predicting Minority Representation: Application to Redistricting and Voting Rights Cases." *American Political Science Review* 115 (4): 1210–1225.
- Axelrod, Robert. 1972. "Where the votes come from: An analysis of electoral coalitions, 1952–1968." *American Political Science Review* 66 (1): 11–20.
- Barreto, Matt A., Loren Collingwood, and Sylvia Manzano. 2010. "A new measure of group influence in presidential elections: Assessing Latino influence in 2008." *Political Research Quarterly* 63 (4): 908–921.
- Barreto, Matt A., Lorrie Frasure-Yokley, Edward D. Vargas, and Janelle Wong. 2018. "Best practices in collecting online data with Asian, Black, Latino, and White respondents: evidence from the 2016 Collaborative Multiracial Post-election Survey." *Politics, Groups, and Identities* 6 (1): 171–180.
- de Benedictis Kessner, Justin. 2015. "Evidence in Voting Rights Act Litigation: Producing Accurate Estimates of Racial Voting Patterns." *Election Law Journal* 14 (4): 361–381.

- Broockman, David E., and Christopher Skovron. 2018. "Bias in perceptions of public opinion among American political elites." *American Political Science Review* 112 (3): 542–563.
- Bürkner, Paul-Christian. 2017. "brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80 (1): 1–28.
- Buttice, Matthew K., and Benjamin Highton. 2013. "How does multilevel regression and poststratification perform with conventional national surveys?" *Political Analysis* 21 (4): 449–467.
- Cain, Bruce E., and Emily R. Zhang. 2016. "Blurred Lines: Conjoined Polarization and Voting Rights." *Ohio State Law Journal* 77:867–904.
- Charles, Guy-Uriel E., and Luis Fuentes-Rohwer. 2014. "The Voting Rights Act in Winter: The Death of a Superstatue." *Iowa Law Review* 100:1389–1439.
- Cho, Wendy K. Tam. 1995. "Asians — A monolithic voting bloc?" *Political Behavior* 17 (2): 223–249.
- . 1998. "If the assumption fits: A comment on the King ecological inference solution." *Political Analysis* 7:143–163.
- Cohn, Nate. 2016. "There Are More White Voters Than People Think. That's Good News for Trump." *New York Times*, June 6, 2016, <https://perma.cc/FHS8-ELFW>.
- Collingwood, Loren, Kassra Oskooii, Sergio Garcia-Rios, and Matt Barreto. 2016. "eiCompare: Comparing Ecological Inference Estimates across EI and EI: R x C." *The R Journal* 8 (2): 92.
- Cramer, Katherine J. 2016. *The Politics of Resentment: Rural consciousness in Wisconsin and the rise of Scott Walker*. Chicago, IL: University of Chicago Press.
- Daily Kos. 2021. "Calculating Statewide Election Results by Congressional and Legislative District: Statement of Methodology," <https://perma.cc/WV9E-KRTV>.
- Davidson, Chandler, and Bernard Grofman. 1994. *Quiet Revolution in the South: The Impact of the Voting Rights Act, 1965-1990*. Princeton, NJ: Princeton University Press.
- Dawson, Michael C. 1995. *Behind the Mule: Race and Class in African-American Politics*. Princeton, NJ: Princeton University Press.

- Donovan, Todd. 2010. "Obama and the White Vote." *Political Research Quarterly* 63 (4): 863–874.
- Elmendorf, Christopher S., and Douglas M. Spencer. 2015. "Administering Section 2 of the Voting Rights Act After *Shelby County*." *Columbia Law Review* 115:2143–2218.
- Enos, Ryan D. 2017. *The Space Between Us: Social Geography and Politics*. Cambridge, UK: Cambridge University Press.
- Erikson, Robert S. 2010. "Hispanic Voting in the American States: The Case of 2004," edited by Rodolfo O. de La Garza, Louis DeSipio, and David L. Leal. Notre Dame, IN: University of Notre Dame Press.
- Erikson, Robert S., Gerald C. Wright, and John P. McIver. 1989. "Political parties, public opinion, and state policy in the United States." *American Political Science Review* 83 (3): 729–750.
- Frasure-Yokley, Lorrie, Janelle Wong, Edward Vargas, and Matt Barreto. 2020. "The Collaborative Multiracial Post-Election Survey (CMPS): Building the academic pipeline through data access, publication, and networking opportunities." *PS: Political Science & Politics* 53 (1): 150–151.
- Freedman, David A., Stephen P. Klein, Jerome Sacks, Charles A. Smyth, and Charles G. Everett. 1991. "Ecological regression and voting rights." *Evaluation Review* 15 (6): 673–711.
- Gay, Claudine. 2001. "The Effect of Black Congressional Representation on Political Participation." *American Political Science Review* 95 (3): 589–602.
- Gelman, Andrew. 2009. *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*. Princeton, NJ: Princeton University Press.
- Gelman, Andrew, and Jennifer Hill. 2006. *Data analysis using regression and multilevel / hierarchical models*. Cambridge, UK: Cambridge University Press.
- Gelman, Andrew, and Thomas C. Little. 1997. "Poststratification into many categories using hierarchical logistic regression." *Survey Methodology* 23 (2): 127–135.
- Gelman, Andrew, et al. 2020. "Bayesian Workflow." *arXiv preprint arXiv:2011.01808*.

- Ghitza, Yair, and Andrew Gelman. 2013. "Deep interactions with MRP: Election turnout and voting patterns among small electoral subgroups." *American Journal of Political Science* 57 (3): 762–776.
- . 2020. "Voter Registration Databases and MRP: Toward the Use of Large-Scale Databases in Public Opinion Research." *Political Analysis* 28 (4): 507–531.
- Ghitza, Yair, and Mark Steitz. 2020. "DEEP-MAPS Model of the Labor Force." *Working Paper*.
- Gimpel, James G., Nathan Lovin, Bryant Moy, and Andrew Reeves. 2020. "The Urban–Rural Gulf in American Political Behavior." *Political Behavior* 42 (4): 1343–1368.
- Grimmer, Justin, Eitan Hersh, Marc Meredith, Jonathan Mummolo, and Clayton Nall. 2018. "Obstacles to estimating voter ID laws' effect on turnout." *The Journal of Politics* 80 (3): 1045–1051.
- Grofman, Bernard. 1991. "Statistics without Substance: A Critique of Freedman et al. and Clark and Morrison." *Evaluation Review* 15 (6): 746–769.
- Grofman, Bernard, Lisa Handley, and David Lublin. 2000. "Drawing effective minority districts: A conceptual framework and some empirical evidence." *North Carolina Law Review* 79:1383–1430.
- Hertel-Fernandez, Alexander, Matto Mildemberger, and Leah C. Stokes. 2019. "Legislative staff and representation in Congress." *American Political Science Review* 113 (1): 1–18.
- Hopkins, Daniel J. 2010. "Politicized Places: Explaining Where and When Immigrants Provoke Local Opposition." *American Political Science Review* 104 (1): 40–60.
- . 2018. *The Increasingly United States: How and why American Political Behavior Nationalized*. Chicago, IL: University of Chicago Press.
- Ingham, Sean. 2019. *Rule by Multiple Majorities: A New Theory of Popular Control*. Cambridge, UK: Cambridge University Press.
- Jacobson, Gary C. 2015. "It's Nothing Personal: The Decline of the Incumbency Advantage in US House Elections." *The Journal of Politics* 77 (3): 861–873.
- Jefferson, Hakeem. 2023. "The Curious Case of Black Conservatives: Assessing the Validity of the Liberal-Conservative Scale among Black Americans." *Public Opinion Quarterly*, *Forthcoming*.

- Kastellec, Jonathan P., Jeffrey R. Lax, Michael Malecki, and Justin H. Phillips. 2015. "Polarizing the electoral connection: Partisan representation in Supreme Court confirmation politics." *The Journal of Politics* 77 (3): 787–804.
- Key, V. O. 1948. *Southern Politics in State and Nation*. Knoxville, TN: University of Tennessee Press.
- Kinder, Donald R., and Lynn M. Sanders. 1996. *Divided by Color: Racial Politics and Democratic Ideals*. Chicago, IL: University of Chicago Press.
- King, Gary. 1996. "Why Context Should Not Count." *Political Geography* 15 (2): 159–164.
- Kousser, J. Morgan. 1974. *The Shaping of Southern Politics: Suffrage Restriction and the Establishment of the One-Party South, 1880-1910*. New Haven, CT: Yale University Press.
- Kuriwaki, Shiro. 2021. "The Swing Voter Paradox: Electoral Politics in a Nationalized Era." PhD diss., Harvard University.
- Kuriwaki, Shiro, Stephen Ansolabehere, Angelo Dagonel, and Soichiro Yamauchi. 2023. "Vote Choice and Population Size by Geography, Demographics, and Turnout." *Harvard Dataverse*.
- Kuziemko, Ilyana, and Ebonya Washington. 2018. "Why did the Democrats lose the South? Bringing new data to an old debate." *American Economic Review* 108 (10): 2830–2867.
- Lauderdale, Benjamin E., Delia Bailey, Jack Blumenau, and Douglas Rivers. 2020. "Model-based pre-election polling for national and sub-national outcomes in the US and UK." *International Journal of Forecasting* 36 (2): 399–413.
- Lax, Jeffrey R., and Justin H. Phillips. 2009. "How Should We Estimate Public Opinion in The States?" *American Journal of Political Science* 53 (1): 107–121.
- Leemann, Lucas, and Fabio Wasserfallen. 2017. "Extending the Use and Prediction Precision of Subnational Public Opinion Estimation." *American Journal of Political Science* 61 (4): 1003–1022.
- Masuoka, Natalie, Hahrie Han, Vivien Leung, and Bang Quan Zheng. 2018. "Understanding the Asian American vote in the 2016 election." *Journal of Race, Ethnicity, and Politics* 3 (1): 189–215.
- McCartan, Cory, et al. 2022. *ALARM Project: 50-State Redistricting Simulations*. <https://doi.org/10.7910/DVN/SLCD3E>.

- McKee, Seth C. 2008. "Rural voters and the polarization of American presidential elections." *PS: Political Science & Politics* 41 (1): 101–108.
- McKee, Seth C., and Melanie J. Springer. 2015. "A tale of "Two Souths": White voting behavior in contemporary Southern elections." *Social Science Quarterly* 96 (2): 588–607.
- Morris, Mitzi, et al. 2019. "Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in Stan." *Spatial and Spatio-temporal Epidemiology* 31:100301.
- Munis, B. Kal. 2020. "Divided by place: The enduring geographical fault lines of American politics." PhD diss., University of Virginia.
- Pildes, Richard H. 2001. "Is Voting-Rights Law Now at War with Itself? Social Science and Voting Rights in the 2000s." *North Carolina Law Review* 80:1517–1574.
- Rodden, Jonathan A. 2019. *Why Cities Lose: The Deep Roots of the Urban-Rural Political Divide*. New York, NY: Basic Books.
- Rosenman, Evan T.R., Cory McCartan, and Santiago Olivella. 2023. "Recalibration of Predicted Probabilities Using the 'Logit Shift': Why does it work, and when can it be expected to work well?" *Political Analysis*.
- Shaw, Daron R. 1997. "Estimating racially polarized voting: A view from the states." *Political Research Quarterly* 50 (1): 49–74.
- Si, Yajuan. 2021. "On the Use of Auxiliary Variables in Multilevel Regression and Poststratification." *arXiv preprint arXiv:2011.00360*.
- Smith, Candis Watts, Rebecca J. Kreitzer, and Feiya Suo. 2020. "The dynamics of racial resentment across the 50 US states." *Perspectives on Politics* 18 (2): 527–538.
- Snyder, James M. 1989. "Political geography and interest-group power." *Social Choice and Welfare* 6 (2): 103–125.
- Stan Development Team. 2021. "Stan Modeling Language Users Guide and Reference Manual 2.28," <https://mc-stan.org>.
- Stephanopoulos, Nicholas O. 2016. "Race, Place, and Power." *Stanford Law Review* 68:1323–1408.
- Voting and Election Science Team. 2020. *2020 Precinct-Level Election Results*. V. 35. <https://doi.org/10.7910/DVN/K7760H>.

- Warshaw, Christopher, and Jonathan A. Rodden. 2012. "How should we measure district-level public opinion on individual issues?" *The Journal of Politics* 74 (1): 203–219.
- White, Ismail K., and Chryl N. Laird. 2020. *Steadfast Democrats: How Social Forces Shape Black Political Behavior*. Princeton, NJ: Princeton University Press.

Online Appendix for Kuriwaki et al., “The Geography of Racially Polarized Voting: Calibrating Surveys at the District Level”

A MODELING METHODOLOGY

A.1 Notation

Throughout the paper and supplemental material, we use the following notation.

Indices:

i	Individuals (up to N , in set \mathcal{I})
g	Racial Group (up to G , in set \mathcal{G})
j	Congressional Districts (up to J , in set \mathcal{J})
s	Cells for demographic variables \times geography (in set \mathcal{S})
m	MCMC posterior samples (up to M)

Variables and Constants:

N	Population size
n	Sample size of respondents
Y	Vote for Republican Presidential candidate (binary)
X, Z	Covariates

Quantities of Interest:

τ	The Republican voteshare in a given geography or race
π	Estimated probability at the cell-level, from a logit regression of Y
δ	A correction factor on the logit scale to better estimate τ
κ	Fraction of the variation in τ explained by components

Parameters:

α	Random effect intercepts
β	Coefficients on demographic variables, random slopes
σ^2, ν^2	Variance governing random effects

A.2 About the Cooperative Congressional Election Study

We use the CCES because it is the only survey dataset whose microdata is publicly available, has sufficient coverage of all 435 congressional districts, measures validated vote, and includes geographic indicators for those districts in its public use file. The CMPS, a multi-lingual survey, imposes a three-year lag on the public release of the data. The face-to-face sample of the American National Election Study (ANES) cannot sample all districts. In its worst year, only one percent of the ANES sample came from a competitive open-seat congressional district (Stoker and Bowers 2002). Exit polls have selection issues in the opposite direction: they oversample battleground districts with more media interest. That explains why earlier survey studies of vote choice by race have only produced state-level estimates.

Like most modern internet surveys, the CCES is partly on an opt-in sample, so it provides reliable inferences only with appropriate adjustments after data collection, such as calibration and post-stratification (Rivers 2007). The CCES takes a larger pool of respondents from an online panel, and then prunes the respondents so that it matches the demographics of the adult population at the state level. Its poststratification weights correct for any remaining imbalances at the state level. The estimated vote for statewide elections is about 2-3 percentage points in root mean square error (Ansolabehere and Rivers 2013).

The Presidential vote question is worded:

```
``For whom did you vote for President of the United States?''
```

- Donald Trump (Republican)
- Jill Stein (Green)
- Hillary Clinton (Democrat)
- Evan McMullin (Independent)
- Gary Johnson (Libertarian)

Table A.1 specifies how we bin and categorize the demographic variables in the CCES. The race and ethnicity question, in particular, is worded in the following way:

```
``What racial or ethnic group best describes you?''
```

- White
- Native American
- Black or African-American
- Middle Eastern
- Hispanic or Latino
- Mixed Race
- Asian or Asian-American
- Other

And for those who do not respond “Hispanic or Latino”, the CCES asks a follow-up question,

``Are you of Spanish, Latino, or Hispanic origin or descent?''

- Yes
- No

As explained in the main text, if a respondent answers Hispanic on either of the two questions, we label that respondent as Hispanic. Non-Hispanics are then coded as Whites, Blacks, or Others according to their responses in the first question. The “Other” category encompasses all categories other than White, Black, and Hispanic.

Table A.1 – Demographic covariates for vote-choice model. We use race, education, age group, and sex as well as Trump’s voteshare to predict the vote choice.

Race	Education	Age	Sex
White	High School or Less	18-24	Female
Black	Some College	25-34	Male
Hispanic	4-Year College	35 - 44	
Other	Post-graduate	45 - 64	
		65 +	

A.3 Implementation of Hierarchical Regression

Functional form The model specification we described in Section translates to, in R notation,

```
Y = (1 + race * educ + age + sex | division / state / cd) +  
  race + pct_white + race_B:cd_pct_B + race_H:cd_pct_H + race_0:cd_pct_0 +  
  s(pct_trump)
```

as a logit regression, where

- Y is a binary variable indicating 1 if the respondent voted for Donald Trump and 0 if they voted for Clinton,
- (... | division / state / cd) indicates there are random effects by every CD, which are nested within states, which are in turn nested within Divisions, with the notation A/B being shorthand for A + A*B,
- (1 + race + ... | ...) indicates there are varying coefficients on race for each of the random effect intercepts on the right hand side of the bar,

- `race` indicates a fixed effect for a national difference between racial groups (as opposed to those that varying coefficients by geography), and is an individual level categorical variable (taking on four levels, with White being the baseline),
- `race_B:cd_pct_B` indicates that an indicator variable for whether the respondent is Black (`race_B`) is interacted with the estimated percent of that individual's district that is Black (`cd_pct_B`), with H indicating Hispanics and 0 indicating the Other racial group, and
- `s(pct_trump, ...)` indicates a flexible spline of Trump's voteshare in each congressional district. Daily Kos estimates this information from precinct results assigning them to their congressional district, and thus this quantity is known almost exactly (Daily Kos 2021).

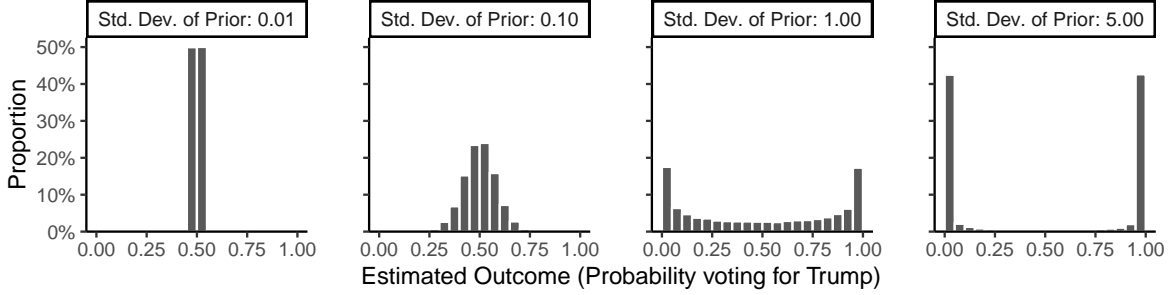
Prior Specification Counter to the intuition that less informative priors imply more flexible values, Figure A.1 shows that less informative priors actually imply more opinionated predicted outcomes (Gelman et al. 2020). To obtain a prior predictive distribution, we sampled only from the prior with the data matrix structure following the terms in the model. A $\text{Normal}(0, 5^2)$ prior for all random effects and intercepts, typically considered an uninformative prior, has enough of a heavy tail that compounds into the large absolute values on the logit scale when combined in dozens of linear combinations. That implies a separation into 0 and 1 on the probability scale shown in the figure, which we believe is too extreme a prior.

In our specifications, we choose a $\text{Normal}(0, 1^2)$ prior. This prior creates enough mass at all values of the support of the outcome (Figure A.1). It has two modes at 0 and 1, but we are willing to accept this as saying that some demographic cells may have a high probability of voting Republican or Democrat.

For the correlation matrix between the geographic random effects, we used an LKJ correlation matrix with a parameter of 1 (default in the package), similar to a flat prior in other classes of priors.

Sampling We estimate samples with 4 chains, using the Hamilton Monte Carlo in Stan. Each chain contains 1000 burn-in samples which are discarded followed by 1000 samples that are used. These 1000 samples are then thinned by 2 to result in 500 samples. Across 4 chains, we are left with 2000 retained samples.

Figure A.1 – Implication of Different Prior Specifications on Prior Predictive Outcome. A fraction of the predictive distribution sampled only from the input prior distribution. Input priors vary in the standard deviation of the Normal and facets are arranged from a tight prior of Normal(0, 0.01) to a less informative prior of Normal(0, 5).



A.4 Connections with Ecological Inference Estimators

Here we show that our survey-based approach can reduce to the same estimator as a ecological inference (EI) estimator with linear contextual effects. Linear contextual effects are difficult to model in EI, but easier in individual-level data in surveys. This section also serves as a more careful explanation of how partial pooling works in our main model.

In the following, let j index districts. For simplicity, we assume that each district has two groups, $g \in \{B, W\}$ (Black voters and White voters). The discussion below applies to the case of multiple groups with additional notation.

Survey Estimator Small area estimation (SAE) uses the survey data and constructs estimators by partially pooling information across areas to reduce variance. We first define the direct estimator for the Black in district j as

$$\hat{\tau}_{Bj}^{\text{dir}} = \frac{1}{n_{Bj}} \sum_{i \in S_{Bj}} Y_i$$

where Y_i is the survey outcome for unit i , S_{Bj} denotes the set of sampled respondents in district j who belong to group B , and n_{Bj} is the sample size of this group. We define the direct estimator for the White $\hat{\tau}_{Wj}$ similarly. Under the standard assumptions, the direct estimator is unbiased and approximately normal $\hat{\tau}_{gj} \sim \mathcal{N}(0, v_g^2/n_{gj})$. This estimator is prohibitively high-variance because typically n_{gj} is small.

The regression-based SAE estimator in its simplest form can be characterized as coming

from the following random effects equation

$$\tau_{gj} = \gamma_g + \delta_g X_j + \epsilon_{gj}, \quad \epsilon_{gj} \sim \mathcal{N}(0, \sigma_g^2) \quad (5)$$

where ϵ is a error term, γ_g is a fixed effect for race and X_j is the proportion Black. The variable X can include other district level variables, but we focus on the proportion Black for a simpler comparison with EI. Note that the main model we estimate in our article is more complex than this, by using varying slopes to effectively estimate demographic subgroups (e.g. individual education and gender) that comprise the population subgroup \mathcal{S}_{Bj} .

The pooling model in Equation (5) models the heterogeneity of the voting behavior of each group as a function of the proportion Black. If the voting behavior of the black voters is uncorrelated with the racial composition, we would have $\delta_B = 0$. As we discuss below, the *ecological inference* model does require that the proportion Black and the voting behavior be uncorrelated at the precinct level, and it must hold for the both racial groups simultaneously $\delta_B = \delta_W = 0$. In contrast, the modeling approach for SAE in Equation (5) allows for $\delta_B \neq 0$. This implies that if the White voters in a majority-White district and the White voters in a minority-White district have different voting behavior, the model is able to capture the heterogeneity as long as it is a function of X_j included in the model.

Furthermore, the formula in (5) clarifies that two groups B and W have different sets of coefficients $\{\gamma_B, \delta_B\}$ and $\{\gamma_W, \delta_W\}$. This implies that heterogeneity of the Black voters and that of the White voters are allowed to be different. For example, Black voters in a majority-White district might behave differently from Black voters in a minority-majority district, while White voters might be more homogeneous across districts. In such a case, we would have a large value of δ_B so that τ_{Bj} varies across districts, while a slope for the White voters might be small δ_W to reflect their homogeneous behaviors.

The estimator of $\hat{\tau}_{gj}^{\text{SA}}$ is a partial pooling estimator, which takes the form

$$\begin{aligned} \hat{\tau}_{gj}^{\text{SA}} &= (1 - w_{gj}) \hat{\tau}_{gj}^{\text{dir}} + w_{gj} (\hat{\gamma}_g + \hat{\delta}_g X_j) \\ &= \hat{\tau}_{gj}^{\text{dir}} + w_{gj} [(\hat{\gamma}_g + \hat{\delta}_g X_j) - \hat{\tau}_{gj}^{\text{dir}}] \end{aligned} \quad (6)$$

where $w_{gj} = \sigma_g^{-2} / (\sigma_g^{-2} + \nu_{gj}^{-2} n_{gj})$ is the relative weight placed on the indirect estimator (i.e., away from the direct estimator). We therefore see that the partial pooling places more weight on the direct estimator as its sample size n_{gj} increases, and places less weight as the variance of the random effect σ_g^2 increases.

The coefficients γ_g and δ_g are estimated by regressing $\widehat{\tau}_{gj}^{\text{dir}}$ on X_j ,

$$\min_{\gamma_g, \delta_g} \sum_{j=1}^J [\widehat{\tau}_{gj}^{\text{dir}} - (\gamma_g + \delta_g X_j)]^2 \quad (7)$$

The formula in Equation (6) shows that the bias of $\widehat{\tau}_{gj}^{\text{SA}}$ is a function of how far the truth τ_{gj} is from the predicted value based on the pooling model,

$$\text{Bias}(\widehat{\tau}_{gj}^{\text{SA}}) = w_{gj} \mathbb{E}[(\widehat{\gamma}_g + \widehat{\delta}_g X_j) - \tau_{gj}] \xrightarrow{n_{gj} \rightarrow \infty} 0. \quad (8)$$

This shows that the bias of the SAE estimate for Black voters in district j tends to be large when the predicted value $\widehat{\gamma}_B + \widehat{\delta}_B X_j$ is far from the truth. For example, the prediction might be inaccurate when the proportion Black X_j explains only a fraction variation of τ_{gj} across districts. This is because with large unexplained variance, two districts with similar characteristics $X_j \approx X_{j'}$ tend to have different voting preferences $\tau_{gj} \neq \tau_{gj'}$. We can reduce the magnitude of the bias by including additional variables at the district level, in addition to the proportion Black. In our main specification of this article we include the Trump vote share at the district level as another covariate in the pooling model, and further include varying slopes for education, sex, and age group for each district random effect. We also note that the specification of the pooling can be flexible and need not be linear. In fact, our main specification includes the Trump vote share via a cubic spline.

In another extreme when $n_{gj} \rightarrow 0$, the SAE approach allows the estimate to reduce to

$$\begin{aligned} \widehat{\tau}_{Wj}^{\text{SA}} &\rightarrow \widehat{\gamma}_W + \widehat{\delta}_W X_j \\ \widehat{\tau}_{Bj}^{\text{SA}} &\rightarrow \widehat{\gamma}_B + \widehat{\delta}_B X_j. \end{aligned} \quad (9)$$

In other words, our model estimates group behaviors such that estimates vary with the contextual variable of racial composition, even with very small samples. As n_{gj} becomes smaller, the small area estimator is pulled towards the national race coefficient γ_g , representing other people of the same race in different areas. But it is also pulled towards δX_j , the group-specific estimate for people who live in areas with a similar racial composition X_j .

EI Estimator On the other hand, consider an ecological inference (EI) for the same quantity of interest, τ_{gj} with $g \in \{B, W\}$. The EI uses data at the level of precincts, which we denote

$h \in \{1, \dots, H\}$. We start from the accounting identity:

$$\tau_{jh} = \tau_{Bjh}X_{jh} + \tau_{Wjh}(1 - X_{jh})$$

where τ_{jh} is the aggregate vote share at precinct h in district j , X_{jh} is the proportion Black of the precinct, and τ_{Bjh} is the unobserved race-specific vote share at the precinct level for Black voters.

The simple form of ecological regression relies on what is known as the constancy equation. Goodman showed that for the EI estimators $\widehat{\tau}_{Wj}^{\text{EI}}$ and $\widehat{\tau}_{Bj}^{\text{EI}}$ to be unbiased, the precinct level race-specific vote share and the proportion Black need to be uncorrelated, $\text{Cov}(\tau_{Bjh}, X_{jh}) = \text{Cov}(\tau_{Wjh}, X_{jh}) = 0$ which assumes that $\text{Cov}(\tau_{Bjh}, X_{jh}) = \text{Cov}(\tau_{Wjh}, X_{jh}) = 0$.

This constancy assumption implies that

$$\begin{aligned} \mathbb{E}(\tau_{jh}|X_{jh}) &= \tau_{Bj}X_{jh} + \tau_{Wj}(1 - X_{jh}) \\ &= \tau_{Wj} + (\tau_{Bj} - \tau_{Wj})X_{jh} \end{aligned}$$

where $\tau_{Bj} = \mathbb{E}(\tau_{Bjh})$ and $\tau_{Wj} = \mathbb{E}(\tau_{Wjh})$.

Under these conditions, we can derive the EI estimator $\widehat{\tau}_{gj}^{\text{EI}}$ with a least squares regression:

$$\min_{\alpha, \beta} \sum_{h=1}^H [\tau_{jh} - (\alpha + \beta X_{jh})]^2, \quad (10)$$

and we obtain

$$\widehat{\tau}_{Wj}^{\text{EI}} = \widehat{\alpha}, \quad \text{and} \quad \widehat{\tau}_{Bj}^{\text{EI}} = \widehat{\alpha} + \widehat{\beta}. \quad (11)$$

The constancy assumption made in the above clearly satisfies this condition.

Put another way, estimates will be *biased* if either of the racial groups' heterogeneity is explained by the racial composition. The survey-based small area estimates can account for heterogeneities across racial groups and districts by adjusting district-level covariates and allowing for race-specific coefficients, whereas the traditional EI in its simplest form cannot run with contextual effects and assumes away heterogeneities. Although linear contextual models have been proposed in ecological inference (Przeworski 1974), estimation of these models remain largely intractable because of under-identification (Ansolabehere and Rivers 1995). This problem arises because EI uses only the aggregate (i.e., not race specific) outcome. The SAE approach avoids the problem by utilizing the race specific direct estimators which are only available in the survey.

A.5 Population Statistics for Target Estimation

First, the population demographics come from the American Community Survey. We obtain the estimated adult population for each of the 160 cells implied by the crossed-combination of the demographics in the CCES, for each of the 435 congressional districts. For 2016, the [age \times sex \times education \times CD] distribution uses 2016 1-year ACS estimates. The [race \times CD] distribution uses the 2014-2018 5-year ACS estimates. We used the 5-year ACS only for race because the 2016 1-year ACS estimates estimated 0 people of racial minorities in some congressional districts such as the at-large district of Wyoming.

Second, we separate these counts into voters and non-voters, by estimating the probability of turnout for each cell. Therefore, we further partition our 70,000-cell ($[160 \times 435]$) table into a $[160 \times 435 \times 2]$ table and then take only the half that represents the voting population. Our synthetic target updates the distribution of voters so that the turnout rate matches the turnout as a proportion of the voting age population at each congressional district.

A.6 Survey-assisted Synthetic Target Estimation

Here we formalize our procedure for estimating a high dimensional target distribution for poststratification. The methods discussed are implemented in the `synthjoint` package.

Motivation In this subsection, we describe our method for quickly integrating marginal and partial joint distributions with the assistance of an individual-level survey dataset. The general idea is to estimate the conditional distribution in the survey data via *regression*, while constraining the parameters so that a predicted marginal distribution matches the population marginal distribution. We work with multinomial logit regressions because most survey outcomes such are categorical instead of continuous.

Let \mathbf{X} denote a set of variables that we have access to the joint distribution in the population, and let Z denote a variable that we only get to know the marginal distribution in the population. The goal is to estimate the joint distribution $p(\mathbf{X}, Z)$.

Suppose that we have survey sample where we observe the joint distribution $p(\mathbf{X}, Z \mid S = 1)$ where $S = 1$ indicates that the distribution is conditional on the survey sample. Unless the survey data is constructed via random sampling from the population, the joint distribution conditional on $S = 1$ does not match the joint distribution of interest, $p(\mathbf{X}, Z \mid S = 1) \neq p(\mathbf{X}, Z)$.

The idea behind our approach is to estimate the conditional distribution of Z given \mathbf{X}

from the survey and estimate $p(Z | \mathbf{X})$. Because $p(\mathbf{X}, Z) = p(Z | \mathbf{X})p(\mathbf{X})$, we can obtain the target joint distribution by estimating the conditional distribution of Z given \mathbf{X} . In our setting, $p(\mathbf{X})$ is already observed in the population.

If we could assume that $Z \perp\!\!\!\perp S | \mathbf{X}$, we would have that

$$p(Z | \mathbf{X}) = p(Z | \mathbf{X}, S = 1).$$

Thus, estimating the relationship between Z and \mathbf{X} in the survey data will provide an unbiased estimate of the joint distribution $\widehat{p}(Z, \mathbf{X}) = \widehat{p}(Z | \mathbf{X}, S = 1)p(\mathbf{X})$. However, the conditional independence assumption is not appropriate when we want to weight on Z . Conditional independence would imply that accounting for Z in the weighting is unnecessary.

Proposed approach Instead of imposing the conditional independence assumption to obtain the conditional distribution $p(Z | \mathbf{X})$, we propose to find a probability distribution that satisfies the following equality constraint:

$$\int p(Z, \mathbf{X})d\mathbf{X} = \int p_\beta(Z | \mathbf{X}, S = 1)p(\mathbf{X})d\mathbf{X},$$

which implies that marginally the predicted distribution $p_\beta(Z | \mathbf{X}, S = 1)$ matches the marginal population target $p(Z)$. This differs from the approach in Kastellec et al. (2015) and the second proposal in Leemann and Wasserfallen (2017), which both estimate the first term on the right hand side but do not enforce a constraint.

Note that the above constraint does not immediately imply that $p(Z | \mathbf{X}) = p(Z | \mathbf{X}, S = 1)$. However, even when the conditional independence assumption does not hold, the above constraint incorporates the population information.

Finally, we estimate $p(Z | \mathbf{X}, S = 1)$ from the survey data such that the above constraint is satisfied.

Data generating process We consider a case where Z is categorical, so that $p(Z | \mathbf{X}, S = 1)$ can be modeled by the multinomial logit. Let $Z_i \in \{1, \dots, K\}$ denote the “marginal” variable for unit i in the survey sample. Then, the multinomial regression is specified as

$$\Pr(Z_i = k | \mathbf{X}_i, S_i = 1) = \frac{\exp(\beta_k^\top \widetilde{\mathbf{X}}_i)}{\sum_{k'=1}^K \exp(\beta_{k'}^\top \widetilde{\mathbf{X}}_i)}$$

where we set $\beta_1 = \mathbf{0}$ for identification. Here, $\tilde{\mathbf{X}}_i$ includes the intercept as well as interaction terms between variables.

Estimation with exact constraints Now, we could estimate β by maximum likelihood, but we wish to impose the constraint from the observed marginal distribution as discussed above. Specifically, we impose the following

$$\underbrace{\Pr(Z = k)}_{\text{population dist.}} = \mathbb{E}_{\mathbf{X} \sim p(\mathbf{X})} \left\{ \frac{\exp(\beta_k^\top \tilde{\mathbf{X}}_i)}{\sum_{k'=1}^K \exp(\beta_{k'}^\top \tilde{\mathbf{X}}_i)} \right\}$$

where the expectation on the right hand side is over the population distribution of \mathbf{X} .

Therefore, we can estimate β by incorporating the additional moment condition. In practice, we can estimate the parameter by constrained optimization. Let $\mathcal{L}_n(\beta)$ denote the log-likelihood function of the multinomial logit, such that

$$\mathcal{L}_n(\beta) \equiv \log \ell_n(\beta), \quad \ell_n(\beta) = \prod_{i=1}^n \prod_{k=1}^K \left\{ \frac{\exp(\beta_k^\top \tilde{\mathbf{X}}_i)}{\sum_{k'=1}^K \exp(\beta_{k'}^\top \tilde{\mathbf{X}}_i)} \right\}^{\mathbf{1}_{\{Z_i=k\}}}$$

where the product is over respondents in the survey data.

We then obtain the estimate by solving the following constrained optimization:

$$\begin{aligned} & \text{maximize} && \mathcal{L}_n(\beta) \\ & \text{subject to} && \mathbf{g}(\beta) = \mathbf{0} \end{aligned}$$

where

$$g_k(\beta) = \Pr(Z = k) - \mathbb{E}_{\mathbf{X} \sim p(\mathbf{X})} \left\{ \frac{\exp(\beta_k^\top \tilde{\mathbf{X}}_i)}{\sum_{k'=1}^K \exp(\beta_{k'}^\top \tilde{\mathbf{X}}_i)} \right\}$$

Estimation for exact constraints with Polya Gamma augmentation We first show that we can find a solution to the above problem by the EM algorithm. In this paper we will not use this exact algorithm and instead approximate it with an additional layer of optimization, but the general form of the procedure is still useful to outline and will be used in the initialization step of our actual algorithm.

Suppose that we fix parameters β_{-k} , and try to estimate β_k (coefficients for the k th cate-

gory). By the EM algorithm derived in Yamauchi (2021), we have the following M-step,

$$\begin{aligned} & \text{maximize} && Q_k(\beta_k) \\ & \text{subject to} && g_k(\beta_k) = 0 \end{aligned}$$

where the objective function is quadratic in β_k after conditioning on the augmented Polya-Gamma random variable ω .

$$Q_k(\beta_k) = -\frac{1}{2}\beta_k^\top (\mathbf{S} + \Sigma_0^{-1})\beta_k + \beta_k^\top (\tilde{\mathbf{X}}^\top \mathbf{d} + \Sigma_0^{-1} \mu_0)$$

where $\mathbf{S} = \tilde{\mathbf{X}}^\top \text{diag}(\{\omega_{ik}\}_{i=1}^n) \mathbf{X}$ and $d_i = \mathbb{E}[\omega_{ik}] \log \left(\sum_{k'=k} \beta_{k'}^\top \tilde{\mathbf{X}}_i \right) + (Z_{ik} - 1/2)$.

We can obtain the update by considering the Lagrangian

$$L(\beta_k, \lambda) = Q_k(\beta_k) + \lambda g_k(\beta_k).$$

The optimality conditions require we solve for

$$\begin{bmatrix} \frac{\partial}{\partial \beta_k} L(\beta_k, \lambda) \\ \frac{\partial}{\partial \lambda} L(\beta_k, \lambda) \end{bmatrix} = \nabla L(\beta_k, \lambda) = \mathbf{0}$$

where each component of the gradient is

$$\begin{aligned} \frac{\partial}{\partial \beta_k} L(\beta_k, \lambda) &= -(\mathbf{S} + \Sigma_0^{-1})\beta_k + (\tilde{\mathbf{X}}^\top \mathbf{d} + \Sigma_0^{-1} \mu_0) + \lambda \frac{\partial}{\partial \beta_k} g_k(\beta_k), \\ \frac{\partial}{\partial \lambda} L(\beta_k, \lambda) &= Q_k(\beta_k) + g_k(\beta_k). \end{aligned}$$

With these gradient functions, we update the estimate of $\tilde{\beta}_k = (\beta_k, \lambda)$ with the Newton-method, so that at iteration $(t + 1)$,

$$\tilde{\beta}_k^{(t+1)} = \tilde{\beta}_k^{(t)} + \left(\nabla^2 L(\beta_k^{(t)}, \lambda^{(t)}) \right)^{-1} \nabla L(\beta_k^{(t)}, \lambda^{(t)})$$

We then implement the E-step by evaluating $\mathbb{E}[\omega_{ik}]$, which follows directly from the mean of a Polya-Gamma random variable,

$$\mathbb{E}[\omega_{ik}] = \frac{1}{2\hat{\psi}_{ik}} \tanh(\hat{\psi}_{ik}/2)$$

where $\widehat{\psi}_{ik} = \beta_k^\top \widetilde{\mathbf{X}}_i - \log \sum_{k' \neq k} \exp(\beta_{k'}^\top \widetilde{\mathbf{X}}_i)$.

Relaxing the exact constraint The exact constraint $g(\beta) = 0$ can lead to unstable estimates when the population distribution and survey data are quite different. In practice, therefore, we do not use the full EM algorithm described above. We instead run a version relaxing the constraint with a slack parameter ϵ . Instead of setting the constraint function g to 0, we reformulate the question as

$$\begin{aligned} & \text{maximize} && \mathcal{L}_n(\beta) \\ & \text{subject to} && \|g(\beta)\|_1 \leq \epsilon \end{aligned} \tag{12}$$

where the constraint bounds the total variation distance between the predicted marginal distribution of Z and the population distribution. ϵ is therefore the sum of absolute deviations on the probability scale. Because the fraction of categories must sum to 1, this deviation is bounded between 0 and 2.

We solve this optimization problem by entering its Lagrangian in R’s coordinate-wise optimization routine which is common, but with a novel initialization step: As initial values of β , we estimate the multinomial with *no* constraints, using the Polya-Gamma augmentation described in the previous section. The fast estimation of the multinomial model at this stage substantially reduces the time required for convergence in the coordinate-wise optimization.

Estimating the joint probability After estimating the model parameters $\widehat{\beta}$ for the conditional distribution, we obtain the population joint probability as

$$\widehat{\Pr}(Z = k, \mathbf{X} = \mathbf{x}) = \frac{\exp(\widehat{\beta}_k^\top \widetilde{\mathbf{x}})}{\sum_{k'=1}^K \exp(\widehat{\beta}_{k'}^\top \widetilde{\mathbf{x}})} \times \Pr(\mathbf{X} = \mathbf{x}).$$

A.7 Estimation of Joint Demographic Distributions

In the next two subsections, we document how we implement the synthetic target estimation algorithm described above. For a poststratification target of the voting age population by each congressional district, we start with the CCES survey data and ACS summary statistics. Although the ACS reports summary statistics at the congressional district level of U.S. adults, it does not report the joint four-way distribution of age, sex, race, and education for each district. We use the following two partitions of the US adult population in each geographic

unit:

- Age group by sex by race: Census table B01001, for subtables B, C, D, F, G, H, I, using variable codes 7–16, 22–31 in each subtable.
- Age group by sex by education: Census table B15001, for variable codes 4–83, excluding variables 11, 19, 27, 35, 43, 44, 52, 60, 68, 76.

These variable codes are also listed in the `ccesMRPprep` package, under `acscodes_sex_age_educ` and `acscodes_sex_age_race`.

Calibration occurs at the CD-level. We estimate a outcome-balanced multinomial logit predicting 4-way categorization of education, such that the weighted proportion of education in each district matches that of the provided ACS table. The predictive slack in equation (12) was set to $\epsilon = 0.01$.

The predictors of the education were racial group interacted with age group, and an intercept for sex. Levels of the variable are recoded to be consistent with the CCES survey data and the ACS, which is described in the main text. In R notation, this amounts to `educ ~ race * age + female`.

Because not all states and certainly not all districts have enough data points to estimate this model, we use a grouping of states defined in Figure A.2. For each district j , we fit the multinomial logit model using all survey responses from the set of states that includes district j .

A.8 Estimation of a Calibrated Turnout Model

We start with the 4-way demographic table from above and wish to estimate the turnout rate in each covariate set s . We again use the CCES, where turnout is 1 if the voter matches to the Catalist voter file in the state and 0 otherwise. This validated vote variable is a standard variable in the public release of the CCES. We use the outcome-balancing multinomial logit predicting a binary variable for turnout, such that the weighted turnout rate in the CD is equal to the observed turnout as a population of the Voting Age Population in the CD. We use the highest office turnout and VAP statistics reported by Daily Kos (2021). We set the predictive slack in equation (12) to $\epsilon = 0.001$.

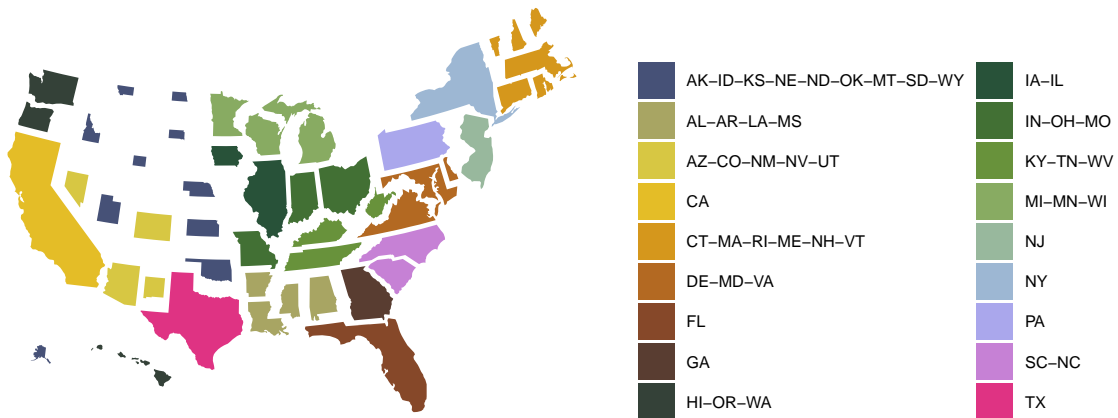
The turnout model takes the form `turnout ~ race * age + female + educ`. Because the contribution of these variables may differ by area, we estimate these models state by state. Because some states' samples have insufficient observations to fit this model, we

again fit the multinomial logit model using all survey responses from the set of states in Figure A.2 that includes the state in question.

A.9 Grouping of States

In survey-assisted synthetic target estimation, it becomes necessary to group small states together or group a small state into a large state so there is sufficient data. We therefore use survey data from the groupings shown in Figure A.2 to estimate district-specific or state-specific targets.

Figure A.2 – Groupings of states for estimates of sufficient sample size.



We grouped states based on geographic proximity and political patterns to group states. Changing the groupings to smaller geographies and simpler regression models changed final estimates by 1-2 percentage points in the average district, for each race.

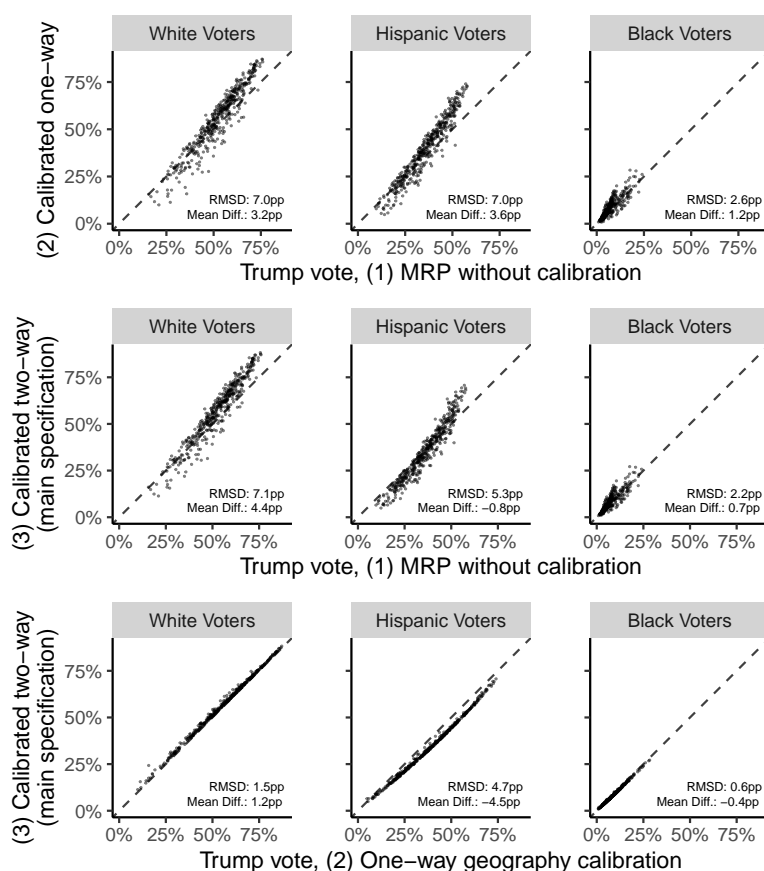
The hierarchical modeling that is used in the main CCES specification can in theory be used to overcome such small sample problems. The reason why we group states in this preparation stage only is because the outcome balancing multinomial model does not partially pool, and limited by the interactions that exist in the data.

A.10 How Modeling Choices Affect Estimates

Figure A.3 shows how our estimates change by the degree of calibration. Here the hierarchical model and the post-stratification are held fixed, but the post-MRP calibration changes.

The graph shows that calibration changes point estimates by around 5-7 percentage points for White and Hispanic voters. Two-way calibration changes the one-way calibration by almost linearly shifting White voters to be more Republican and Hispanic voters to be less Republican in this case. In practice, we implement the two calibrations in the calibration step simultaneously, instead of one calibration after another.

Figure A.3 – Consequences of Calibration. Each row of three scatter plots show the differences in final point estimates depending on the level of calibration discussed in the paper. (1) MRP estimates without any calibration, (2) MRP estimates with a one-way calibration to district-level voteshare, and (3) MRP estimates with a two-way calibration developed in this paper to district-level voteshare and a national vote by race constraint. Statistics show the root mean square difference (RMSD) and mean difference.



B ADDITIONAL ANALYSES

B.1 Asian Americans

In our main analysis, we focus on Black and Hispanic voters and do not separate out Asian American voters from remaining Other racial minorities. The CCES and ACS does record Asian American as a response option in their data and it is possible to construct a hierarchical model that makes this distinction. However, the synthetic joint population estimation currently cannot model separate population counts for Asian Americans because its population is heavily concentrated in a handful of states.

In this section, we analyze the Asian American vote separately but only using data from the states in which a sufficiently large fraction of the population are Asian Americans. Figure B.1 shows point estimates and compares them with the other three racial groups.

Figure B.1 – Comparison of Asian American voting estimates in select states. We estimated joint poststratification tables separating out Asian Americans from the other races, in states with at least 10 percent of Asian American adults. Those states are California, Washington, Hawaii, and New Jersey. Respondents from those four states are used in estimation. Only CDs with an estimated 10 percent or more of the electorate being Asian American are shown.

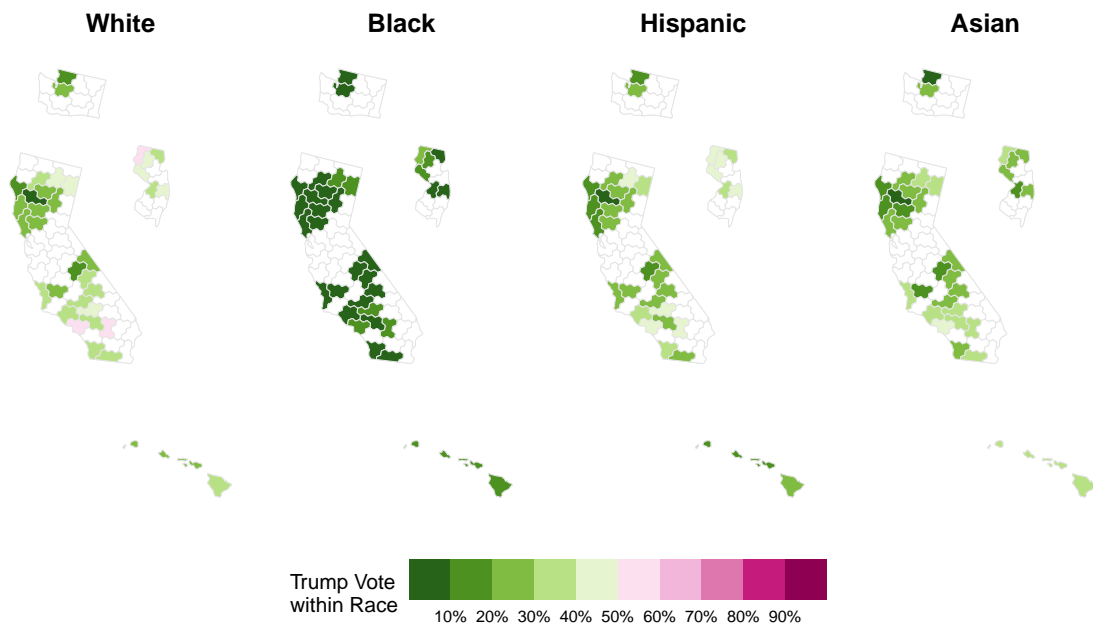


Table B.1 – Comparison of Survey Estimates. All estimates show two party Republican vote in the Presidential election. Exit Polls use the National Election Poll (as published in CNN). *Catalist* refers to Catalist’s MRP estimates. *LD* refers to Latino Decisions and the American Election Eve Poll. *CMPS* refers to the Collaborative Multi-Racial Post-Election Survey. See Appendix C.4 for standard errors for estimates in this paper.

Republican vote among	2016					2020				
	This Paper	Exits	Catalist	LD	CMPS	This Paper	Exits	Catalist	LD	CMPS
National										
White	59	61	59		57	57	59	56	58	52
Black	7	8	7		5	10	12	10	9	10
Hispanic	30	30	29	19	19	38	33	36	28	23
Wisconsin										
White	54	56	54			53	53	53	57	
Black	4	6	8			8	8	10	5	
Hispanic	30	35	34	10		33	38	44	22	
Florida										
White	63	67				62	63		57	
Black	7	9				10	10		9	
Hispanic	39	36		32		47	46		39	
Texas										
White	71	73				66	67		66	
Black	7	12				9	9		10	
Hispanic	38	36		17		43	41		30	

Note: Numbers are taken from the National Exit Polls listed by CNN (2016, 2020), *Catalist website*, Latino Decisions website (2016, 2020), Collaborative Multiracial Post-election Survey (2016, 2020). Blank cells indicate the survey has not publicly released an estimate.

B.2 Comparison with Other Surveys

Table B.1 compares the estimates from other sources of data. These differences could be due to simple random sampling error, differences in turnout estimation, timing of the survey, differences in how racial groups are defined, and survey coverage. We do not have access to the raw data or methodology of all other surveys to conduct a full-fledged comparison.

One discrepancy that stands out is that the estimates by Latino Decisions (LD) and CMPS of the Hispanic vote are more Democratic than the other surveys by double digits. Barreto, Reny, and Wilcox-Archuleta (2017) discussed some reasons for why this might be the case. Discussing the Exit Poll, they point out that a bad selection of precincts to poll might have biased the estimates, and they also question whether the Exit Poll’s composition of the His-

panic respondents is consistent with the Census. These two critiques would apply less to an online survey like the CCES that is then weighted to Census composition figures. A third potential reason is the lack of the Spanish language option in the CCES. A fourth potential reason is that LD and CMPS could be using a different definition of Latinos and recruitment method than the CCES or Exit Polls.

Barreto, Reny, and Wilcox-Archuleta (2017) use precinct-level data as suggestive evidence that the Hispanic Republican vote is much lower than what the Exit Poll has reported. The thrust of such evidence is a homogeneous precinct analysis. Focusing on New York state, for example, Barreto (2016) shows that precincts in which over 80 percent of the electorate are estimated to be Hispanic often reported a Clinton vote of over 90 percent. The implication is that the Exit Poll’s estimates of a 23 percent Trump vote among New York Hispanics is unreasonable. However, our validation analysis in Florida suggests that ecological inference may be underestimating the true Republican vote among Hispanic voters. Because our study provides district level estimates *below* the state-level, we can provide a suggestive test of this independence assumption.

B.3 Details on Florida Validation Analysis

Data We use Catalist’s cleaned voterfile query tool to extract precinct-level aggregates and joint counts of race and party registration. In September 2021, we queried the number of voters (registered in Florida) who voted in the 2016 Presidential Election (wherever they were in 2016). Catalist assigns these voters to the precinct that they are registered in at the time of the query. The race and party registration data are information as of the time of the query as well (not the time of the 2016 election).

After assigning congressional districts to split precincts, we are left with precincts comprising 8.3 million voters. There were 9.4 million votes in the 2016 election. Some of the drop in population is due to the snapshot of the voterfile being pulled not being accurate. We sample 10,000 posterior iterations for each congressional district.

To evaluate the validity of the EI method, we take the actual share of each racial group in each district that are Republican from the same precinct-level dataset. Separately, we purchased a CD-level dataset from Catalist that records the counts of each race and party registration of 2016 voters by their 2016 districts. Because we do not want to attribute errors in EI to discrepancies between our precinct data query and the ground truth, we evaluate the EI estimates based on the ground truth calculated from the same precinct dataset and evaluate the MRP estimates based on the ground truth based on the more accurate counts purchased from

Catalist. The two CD-level ground truth measures are correlated at 0.97 for White voters, 0.98 for Hispanic voters, and 0.89 for Black voters.

Setting Florida and North Carolina are the only two states that ask voter registrants to identify with both a party and a racial group, and makes this information, along with precinct and turnout information, public in their voterfile. Therefore, the proportion of voters in a specific district and of a specific race who register with the Republican party is known precisely through an administrative dataset. We therefore conduct a validation of our MRP method and ecological inference (EI) where the quantity of interest is

$$\Pr(\text{Registration} = \text{Republican} \mid \text{Race } g, \text{ District } j, \text{ Turnout}),$$

that is, the proportion of the electorate in district j with racial identification g that is registered for the Republican party.

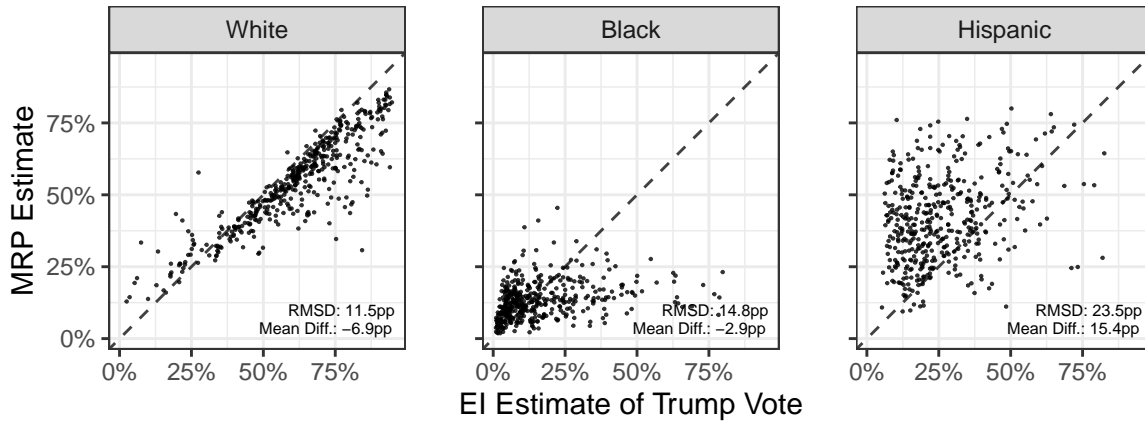
MRP methodology The CCES includes a variable for party registration that comes directly from the Catalist voter file match. In other words, in the CCES party registration is observed without error. We estimate a MRP model predicting party registration in Florida with the same methods described in this paper with a few differences. First, we estimate the model only on Florida data, since not all states have party registration on the voter file. We also evaluated an MRP estimate with only one-way calibration to district geography rather than a two-way calibration. The error rate increases for White voters, but the MRP error rate for Hispanic and Black voters are still lower than that for EI.

Ecological Inference in Practice This setting is a favorable data setting for ecological inference of racial polarized voting than the modal case. The registered race and the party registered race is exactly known in this voter file data. In 48 other states, this data is unobserved. In all but six southern states, analysts use Census estimates the composition of each race in the (*citizen*) *voting age population* as their measure of race at the precinct level, or they use a race classification estimator such as Bayesian-Improves Surname Geocoding to predict race within the voter file. Party registration is only available in 31 states as well.

B.4 Comparison with Ecological Inference in 50 States

We next extend our comparison between our survey-based method and EI to other states, beyond where validation data is available. We use Census and election results data at the

Figure B.2 – Comparison of MRP and Ecological Inference Estimates. Both model predicts the proportion of each racial group electorate in each congressional district that voted for the Republican Presidential candidate in 2020. Statistics in each facet show root mean square difference (RMSD) and mean difference between EI and MRP estimates.

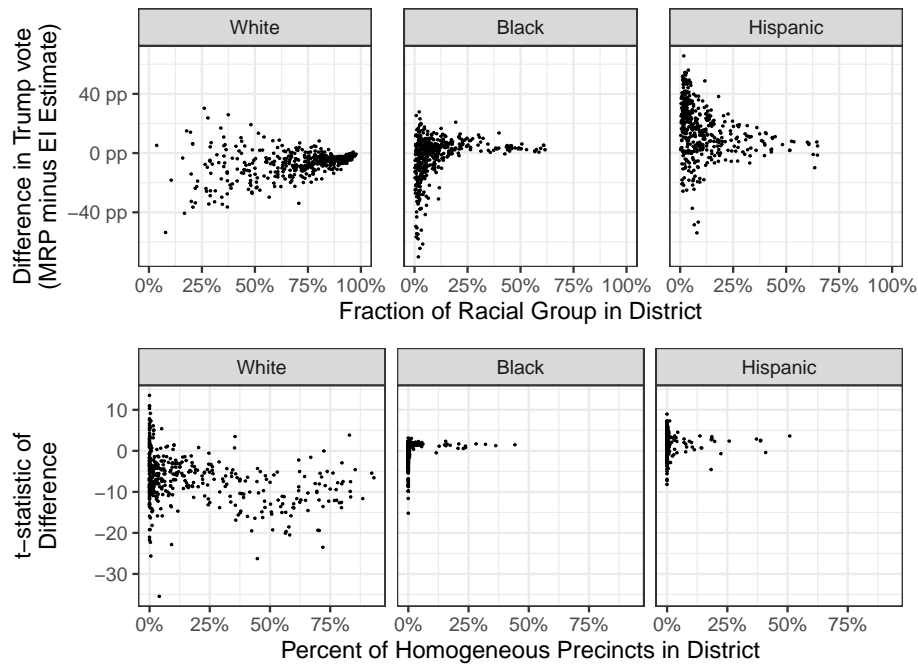


precinct-level which are then assigned to districts as of 2020. We use the combined file from McCartan et al. (2022), which incorporates election data and shapefiles from Voting and Election Science Team (2020) with Census demographics. We compare our method's and EI's estimate for the 2020 election instead of the 2016 election, because the 2020 election is done closer to the 2020 decennial Census that provided the racial composition information used for EI. Our comparison covers 435 districts in all 50 states.

We take racial compositions within the Voting Age Population (VAP). This is different from the Citizen Voting Age Population or the turnout population, but the decennial Census does not ask citizenship and standard ecological inference must rely on such incomplete data. Within each congressional district, we take 5000 thinned draws from the 4 by 2 ecological inference estimates, estimated the same way as Appendix B.3.

Figure B.2 compares estimates of Trump vote share for White, Black and Hispanic voters in a congressional district, with EI estimates on the x-axis and our 2020 MRP estimates on the y-axis. Among White voters, MRP often produces lower estimates of Republican voteshare than EI with an average discrepancy of 12 percentage points in root mean square difference. Among Hispanic voters, we find the opposite, where MRP often produces higher estimates of Republican voteshare than EI. The discrepancy here is even larger, about 23 percentage points. Hispanic estimates have more uncertainty in them due to small subgroup samples, but even accounting for the standard errors of both the MRP and EI estimates, the difference

Figure B.3 – MRP and EI Differences by Group Size and Homogeneity. The top graph plots the raw difference between district-group level MRP and EI estimates on the y-axis, and the district level group size on the x-axis, for White, Black, and Hispanic voting populations. The bottom graph plots the t-statistic for district-group’s MRP and EI estimates on the y-axis, against the percent of homogeneous precincts (precincts with or greater than 90 percent of a group voting population) in that district on the x-axis.



is statistically significant at a 0.01 level in 52 percent of the districts under consideration.

States with large Hispanic voting populations have much smaller differences between MRP and EI estimates of Republican voteshare. The RMSD for Hispanic MRP and EI estimates is between 11 and 16 percentage points in Southwestern states like Arizona, California, and Texas. In states with small Hispanic voting populations like Alabama and Mississippi, the discrepancy is above 40 percentage points.

Similarly, states with larger Black voting populations demonstrate smaller differences between MRP and EI estimates of Republican voteshare. In Southern states like Alabama, Florida, and Georgia, the RMSD for Black MRP and EI estimates is between 6 and 10 percentage points. However in Southwestern states like Arizona, California and Colorado, the RMSD is between 20 and 24 percentage points. Overall, EI estimates Black voters to be 14 percentage points more Republican than MRP estimates.

Figure B.3 plots the differences between MRP and EI estimates against district-group

Table B.2 – MRP and EI White - Non-White Racial Gap. Each column represents summary statistics of MRP and EI reliant estimates of the White - non-White racial gap (Equation (1)) using the point estimates from Figure B.2.

Statistic	MRP	EI
Minimum	0.04	-0.28
District in 25th Percentile	0.14	0.28
Median District	0.22	0.43
Mean district	0.24	0.41
National Gap	0.29	0.42
District in 75th Percentile	0.30	0.57
Maximum	0.65	0.86
Standard deviation	0.12	0.23

level demographics. The top graph demonstrates that differences between MRP and EI estimates decrease towards zero as the size of a group voting population increases. The bottom graph plots the relationship between the t-statistic of the difference measuring the difference between MRP and EI estimates for a district, against a district's percent of homogeneous precincts. Estimates for a group differ most when that group has little to no homogeneous precincts. For Black and Hispanic populations, the difference between MRP and EI estimates decreases as their percentage of homogeneous precincts increases.

In contrast, MRP and EI differences grow larger as the percentage of homogeneously White precincts increases. In these districts with large White populations, an increasing percentage of homogeneously White precinct amounts to an increasingly large number of White voters, which may introduce greater variance in preferences.

The differences between MRP and EI estimates shown in Figure B.2 lead to different estimates of the racial gap. EI estimates White voters to be more cohesively Republican and Hispanic voters to be more cohesively Democratic, leading to larger estimates of the White - non-White racial gap. Table B.2 compares summary statistics on the absolute difference between White and non-White estimates of Trump voteshare. EI produces larger racial gap estimates than MRP across nearly all summary statistics. The median White - non-White racial gap is 22 percentage points under MRP, but 43 percentage points under EI.

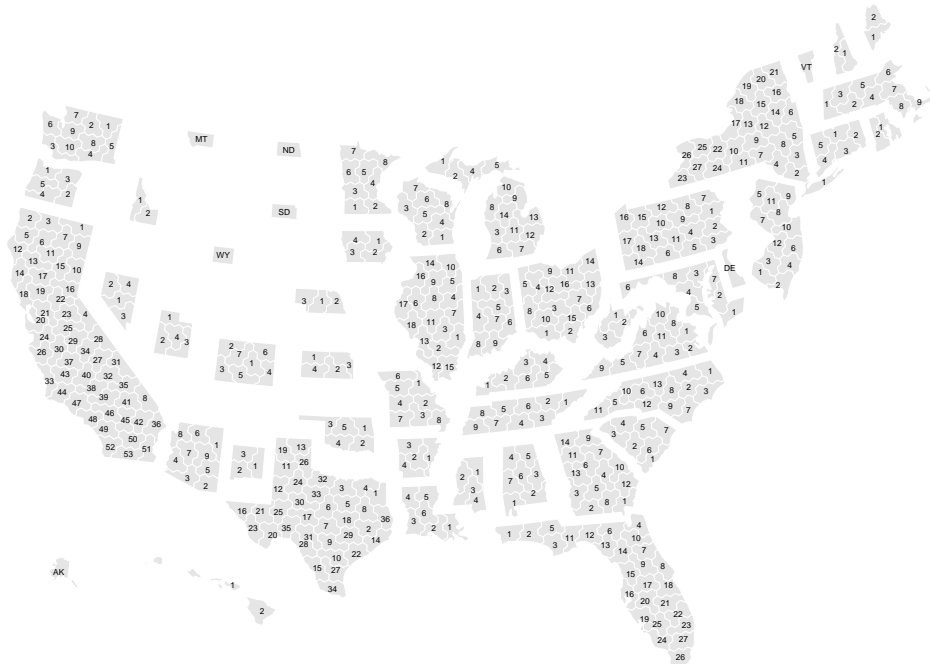
C ADDITIONAL ESTIMATES AND STANDARD ERRORS

C.1 Additional Estimates

A table of estimates for each district and race in 2016 and 2020 are provided in our *APSR* Dataverse repository. A copy of the data is also deposited in a separate Dataverse repository, <https://doi.org/10.7910/DVN/MAZMJ6> (Kuriwaki et al. 2023), for post-stratification tables beyond the scope of years in this study.

The district's names and descriptions are all as of 2021. The names of the CD come from the Daily Kos [name](#) in August 2019 (Daily Kos 2019). When there was significant redistricting between the 2016 and 2020 elections, we model the contemporaneous districts and name them separately. Pennsylvania redrew its Congressional districts in 2018, so we designated districts with the help of Daily Kos and Lara Putnam. The geographic associaton is shown in Figure C.1 and the names of each district.

Figure C.1 – Location of specific Congressional Districts. The Daily Kos district map used throughout this paper places congressional districts roughly according to their relative place in the state. The following figure indicates which districts are placed where.



C.2 2020 Estimates

While this paper focused on 2016 in the main text, we estimated the same quantities using the 2020 Presidential election between Donald Trump and Joe Biden. Weighting and calibration targets were updated to use 2020 statistics. One exception is the voting age population statistics of the ACS. The ACS has only released experimental data for 2020 because of low response rates during COVID-19. We instead take the average of 2019 and 2021.

We took the 2016 and 2020 estimates and compared the estimates of each CD-race combination. That said, these comparisons cannot fully represent the underlying change in a racial group's overall voting preferences because of differential turnout.

Table C.1 shows the results of these estimates at the state-level and up. The format mirrors Table 1 in 2016.

Figure C.2 shows that Whites and Non-Whites moved in opposite directions, in a direction that reduced racial polarization. Whites became less supportive of the Republican Presidential candidate by about 2.5 percentage points and Non-Whites became more supportive by about 3.5 points. Hispanics moved more than Blacks but both moved in the same direction. In other words, the difference in Republican vote between White voters and non-White voters decreased by over 6 percentage points.

C.3 Racial Gap Estimates

Figure C.3 and Table C.2 show the full distribution and summary statistics of the racial gap measure discussed in the main text.

Figure C.4 separates the White - non-White racial gap into the White - Black racial gap and the White - Hispanic racial gap.

Table C.1 – 2020 Republican Vote by Race and Geography

(a) Region Level

	Non-Whites				
	White	All	Black	Hisp.	All
Northeast	48	22	10	29	42
South	66	27	10	44	53
North Central	56	24	9	37	51
West	48	31	12	34	42
National	56	28	10	38	48

(b) Division Level

	Non-Whites				
	White	All	Black	Hisp.	All
Northeast					
New England	40	25	12	28	38
Middle Atlantic	52	22	9	29	44
South					
South Atlantic	62	25	10	43	49
East South Central	73	22	11	60	62
West South Central	69	34	11	44	56
North Central					
East North Central	56	21	8	35	50
West North Central	58	32	14	42	55
West					
Mountain	56	39	16	40	51
Pacific	43	29	10	32	37
National	56	28	10	38	48

(c) State Level

	Non-Whites				
	White	All	Black	Hisp.	All
New England					
Connecticut	44	28	15	34	40
Maine	46	33	(12)	(38)	46
Massachusetts	36	22	10	22	33
New Hampshire	47	36	(16)	(33)	46
Rhode Island	42	28	11	33	40
Vermont	32	22	(8)	(24)	32
Middle Atlantic					
New Jersey	50	27	11	34	42
New York	49	20	9	26	40
Pennsylvania	56	19	8	32	50
South Atlantic					
Delaware	48	22	12	33	41
Florida	62	34	10	46	52
Georgia	70	20	9	45	50
Maryland	46	15	6	25	33
North Carolina	62	24	13	40	51
South Carolina	70	20	10	59	56
Virginia	56	22	9	33	45
West Virginia	72	38	16	(51)	70
East South Central					
Alabama	79	22	12	71	63
Kentucky	68	26	11	(48)	63
Mississippi	82	19	13	(73)	59
Tennessee	70	22	7	54	62
West South Central					
Arkansas	72	30	12	55	64
Louisiana	79	23	13	61	60
Oklahoma	71	47	14	62	67
Texas	65	35	9	42	53
East North Central					
Illinois	50	19	6	32	42
Indiana	62	28	14	42	58
Michigan	55	21	7	40	49
Ohio	59	19	9	(38)	54
Wisconsin	53	27	8	35	50
West North Central					
Iowa	56	37	16	46	54
Kansas	61	41	20	51	58
Minnesota	50	25	11	30	46
Missouri	63	26	12	(44)	58
Nebraska	63	41	20	49	60
North Dakota	69	51	(27)	(53)	67
South Dakota	64	55	(22)	(57)	64
Mountain					
Arizona	57	36	15	35	50
Colorado	47	33	15	35	43
Idaho	67	62	(29)	59	66
Montana	59	54	(30)	(53)	59
Nevada	55	41	16	50	49
New Mexico	51	39	(19)	40	45
Utah	63	51	(27)	48	61
Wyoming	73	66	(34)	67	73
Pacific					
Alaska	58	48	17	46	55
California	42	28	10	31	35
Hawaii	42	31	14	26	35
Oregon	44	32	(9)	33	42
Washington	43	31	10	34	40
National	56	28	10	38	48

Figure C.2 – Changes in Republican vote by racial group, 2016 to 2020. Each point is our CD-level estimate for a racial group. We exclude CDs in Pennsylvania and North Carolina because district lines changed between 2016 and 2020 in those states. Summary statistics show mean change in a pair of CDs without taking the absolute value.

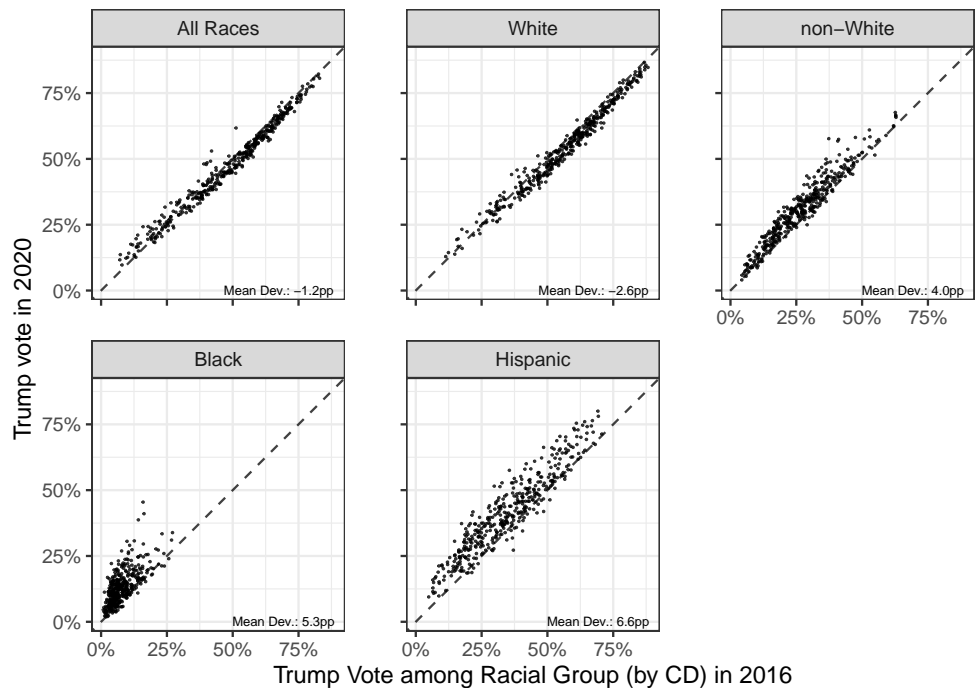


Figure C.3 – Distribution of CD-level Racial Gap Estimates. A histogram of the distribution shown in Figure 3.

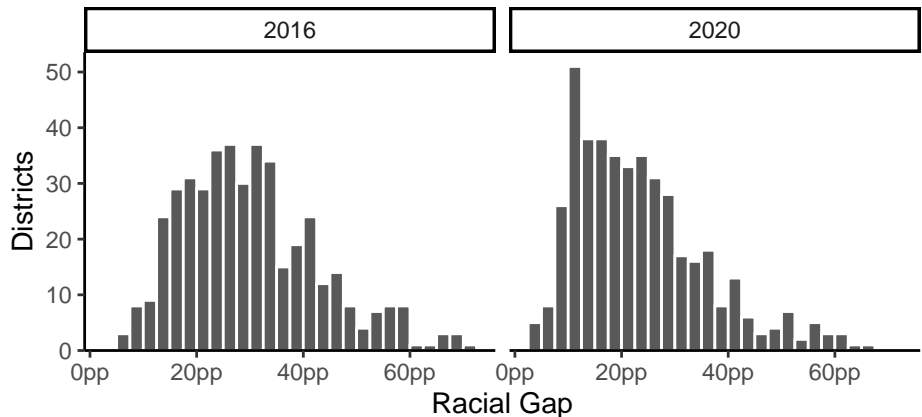
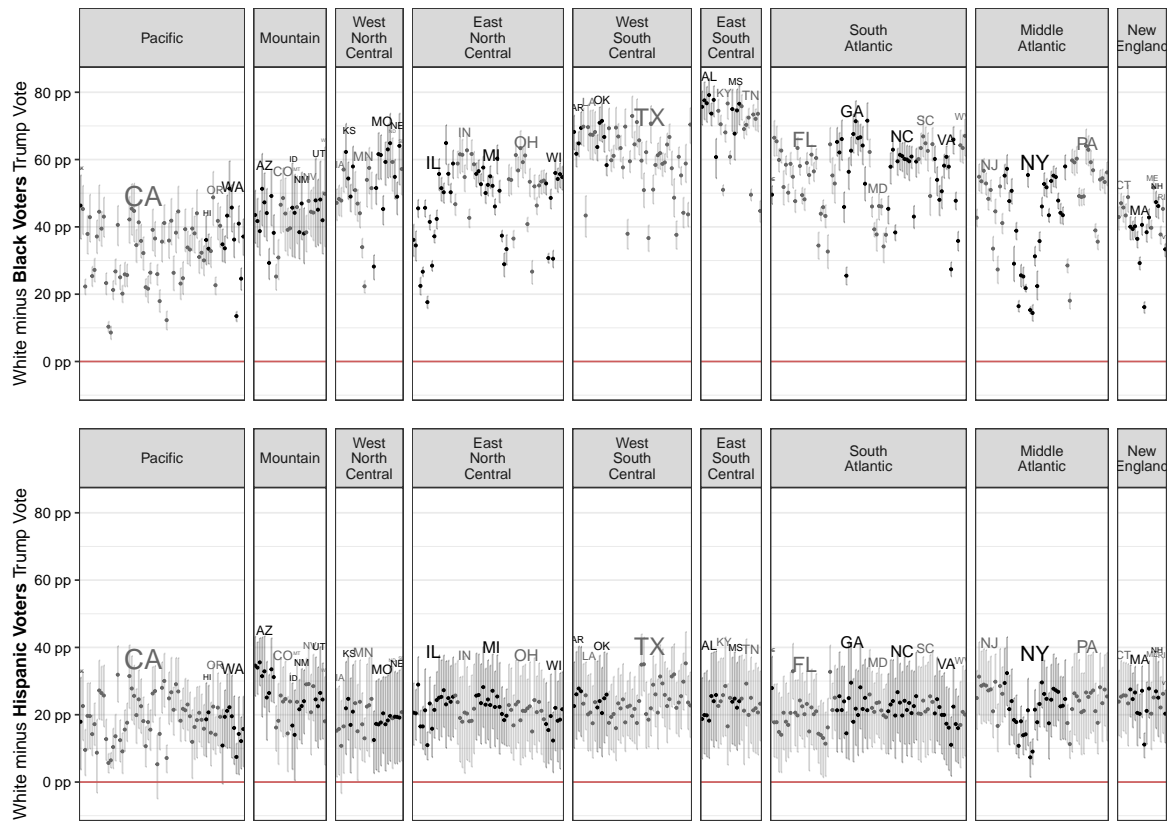


Table C.2 – Summary Statistics of CD-level Racial Gap

Statistic	2016	2020
Minimum	0.06	0.04
District in 25th Percentile	0.21	0.14
Median District	0.29	0.22
Mean District	0.30	0.24
National Gap	0.36	0.29
District in 75th Percentile	0.39	0.30
Maximum	0.71	0.65
Standard Deviation	0.13	0.12

Figure C.4 – The White-Black and White-Hispanic Racial Gap. Bars show 80 per cent credible intervals. The arrangement of districts follows Figure 3.



C.4 Standard Errors

Table 1 compares point estimates, but some of the differences between state-level point estimates are statistically not distinguishable from 0 because of estimation uncertainty. We approximate the state-level standard error by taking the standard deviation of the state-level estimates across the 2000 posterior draws. There are two types of uncertainty that are of interest: the uncertainty implied by the hierarchical regression model, and the uncertainty after calibration is applied.

The standard error implied by the hierarchical model is the standard deviation of the MRP estimates before calibration. It reflects the uncertainty in making inferences from the survey data. These standard errors will likely be large in districts and racial groups with fewer respondents. It is also smaller than the standard error of simple direct estimates without hierarchical model with shrinkage. Table C.3(a) uses the standard deviation of these estimates. A value of 0.05, for example, indicates a frequentist margin of error of about 10 percentage points.

The standard error of the estimates after calibration is important because they are the final estimates we produce. By design, calibration does reduce the variance of estimates of homogeneous racial groups. Consider a state (district) such as VT-01 that is 99 percent White and 1 percent minority. The voteshare in that district is known. The racial homogeneity allows us to attribute the bulk of the calibration shift in the homogeneous racial group, and every posterior iteration of this estimate will be drawn to that voteshare target, dramatically reducing the standard error. In other words, once we know that Trump's voteshare in VT-01 was exactly 34 percent, there is little uncertainty that the White vote in Vermont is also quite close to 34 percent. This idea is consistent with thinking of calibration as a principled posterior update (Rosenman, McCartan, and Olivella 2023). Table C.3(b) uses the standard deviation of these estimates.

Table C.3 – Standard Errors of State-level Estimates. See Appendix C.4 for the difference between the two specifications.

(a) Implied by Hierarchical Model						(b) After Calibration						
		White	Non-Whites				White	Non-Whites				
			All	Black	Hisp.			All	All	Black	Hisp.	
New England						New England						
	Connecticut	.021	.031	.024	.064	.019		Connecticut	.0069	.024	.024	.053
	Maine	.023	.047	.024	.076	.023		Maine	.0015	.043	.025	.07
	Massachusetts	.015	.023	.016	.041	.014		Massachusetts	.0043	.019	.016	.033
	New Hampshire	.023	.043	.026	.075	.023		New Hampshire	.0021	.04	.029	.071
	Rhode Island	.027	.041	.024	.076	.025		Rhode Island	.0065	.033	.023	.063
	Vermont	.03	.044	.019	.068	.03		Vermont	.0043	.039	.019	.057
Middle Atlantic						Middle Atlantic						
	New Jersey	.016	.021	.016	.042	.013		New Jersey	.0077	.015	.016	.033
	New York	.012	.014	.013	.028	.0099		New York	.0053	.01	.012	.021
	Pennsylvania	.011	.014	.013	.046	.01		Pennsylvania	.0026	.012	.012	.039
South Atlantic						South Atlantic						
	Delaware	.026	.024	.022	.08	.022		Delaware	.0067	.018	.019	.068
	Florida	.01	.016	.014	.03	.0092		Florida	.0056	.011	.013	.023
	Georgia	.015	.015	.015	.059	.012		Georgia	.0072	.012	.015	.054
	Maryland	.017	.015	.016	.054	.013		Maryland	.0072	.011	.013	.044
	North Carolina	.015	.015	.015	.053	.012		North Carolina	.0053	.013	.016	.049
	South Carolina	.019	.019	.021	.071	.016		South Carolina	.0066	.017	.02	.07
	Virginia	.014	.017	.015	.05	.012		Virginia	.0056	.013	.014	.043
	West Virginia	.021	.026	.029	.082	.02		West Virginia	.0027	.027	.034	.081
East South Central						East South Central						
	Alabama	.02	.017	.018	.078	.016		Alabama	.008	.019	.023	.072
	Kentucky	.02	.021	.012	.076	.018		Kentucky	.0025	.02	.014	.074
	Mississippi	.023	.022	.024	.085	.017		Mississippi	.012	.026	.03	.082
	Tennessee	.017	.015	.014	.068	.015		Tennessee	.0029	.015	.015	.065
West South Central						West South Central						
	Arkansas	.021	.024	.023	.08	.019		Arkansas	.005	.022	.025	.076
	Louisiana	.019	.024	.028	.069	.016		Louisiana	.012	.021	.028	.066
	Oklahoma	.02	.03	.023	.077	.018		Oklahoma	.0044	.025	.026	.074
	Texas	.012	.019	.014	.029	.011		Texas	.007	.011	.015	.019
East North Central						East North Central						
	Illinois	.013	.015	.014	.034	.011		Illinois	.0045	.012	.013	.028
	Indiana	.016	.021	.018	.059	.015		Indiana	.0027	.02	.021	.058
	Michigan	.014	.016	.014	.05	.013		Michigan	.0029	.015	.015	.047
	Ohio	.012	.014	.015	.051	.011		Ohio	.0023	.015	.016	.05
	Wisconsin	.016	.023	.015	.057	.015		Wisconsin	.0026	.021	.016	.053
West North Central						West North Central						
	Iowa	.019	.04	.031	.08	.018		Iowa	.0031	.037	.035	.078
	Kansas	.021	.038	.03	.079	.02		Kansas	.0047	.034	.036	.076
	Minnesota	.017	.028	.021	.058	.016		Minnesota	.0035	.025	.022	.053
	Missouri	.016	.02	.022	.064	.014		Missouri	.0033	.022	.027	.064
	Nebraska	.022	.041	.032	.083	.021		Nebraska	.0044	.038	.041	.08
	North Dakota	.03	.05	.048	.12	.03		North Dakota	.0045	.042	.063	.12
	South Dakota	.03	.058	.052	.11	.029		South Dakota	.0043	.05	.061	.11
Mountain						Mountain						
	Arizona	.017	.029	.043	.04	.015		Arizona	.0073	.021	.043	.031
	Colorado	.017	.031	.041	.044	.016		Colorado	.0069	.023	.042	.036
	Idaho	.025	.054	.081	.086	.025		Idaho	.0048	.049	.094	.083
	Montana	.029	.056	.074	.11	.028		Montana	.0042	.05	.081	.1
	Nevada	.021	.033	.046	.055	.02		Nevada	.011	.023	.046	.044
	New Mexico	.024	.042	.062	.055	.025		New Mexico	.019	.023	.058	.037
	Utah	.023	.044	.063	.065	.022		Utah	.005	.038	.073	.061
	Wyoming	.031	.064	.087	.1	.031		Wyoming	.012	.058	.11	.1
Pacific						Pacific						
	Alaska	.03	.057	.07	.11	.03		Alaska	.013	.041	.072	.1
	California	.011	.015	.026	.023	.0095		California	.0064	.0072	.021	.013
	Hawaii	.03	.052	.051	.077	.033		Hawaii	.031	.022	.037	.057
	Oregon	.018	.033	.027	.057	.017		Oregon	.004	.027	.025	.049
	Washington	.015	.025	.026	.05	.014		Washington	.004	.02	.024	.043

REFERENCES FOR APPENDIX

- Ansolabehere, Stephen, and Douglas Rivers. 1995. "Bias in Ecological Regression." *MIT Department of Political Science Manuscript*.
- . 2013. "Cooperative Survey Research." *Annual Review of Political Science* 16:307–329.
- Barreto, Matt A. 2016. "2016 Latino Election Analysis (November 30, 2016)," https://latinodecisions.com/wp-content/uploads/2019/06/PostElection2016_-_Barreto_-_CAP.pdf.
- Barreto, Matt A., Tyler Reny, and Bryan Wilcox-Archuleta. 2017. "Survey methodology and the Latina/o vote: Why a bilingual, bicultural, Latino-centered approach matters." *Aztlan: A Journal of Chicano Studies* 42 (2): 211–227.
- Daily Kos. 2019. "The Daily Kos Elections naming guide to the nation's congressional districts," <https://perma.cc/P5C9-DRN8>.
- . 2021. "Calculating Statewide Election Results by Congressional and Legislative District: Statement of Methodology," <https://perma.cc/WV9E-KRTV>.
- Gelman, Andrew, et al. 2020. "Bayesian Workflow." *arXiv preprint arXiv:2011.01808*.
- Kastellec, Jonathan P., Jeffrey R. Lax, Michael Malecki, and Justin H. Phillips. 2015. "Polarizing the electoral connection: Partisan representation in Supreme Court confirmation politics." *The Journal of Politics* 77 (3): 787–804.
- Kuriwaki, Shiro, Stephen Ansolabehere, Angelo Dagonel, and Soichiro Yamauchi. 2023. "Vote Choice and Population Size by Geography, Demographics, and Turnout." *Harvard Dataverse*.
- Leemann, Lucas, and Fabio Wasserfallen. 2017. "Extending the Use and Prediction Precision of Subnational Public Opinion Estimation." *American Journal of Political Science* 61 (4): 1003–1022.
- McCartan, Cory, et al. 2022. *ALARM Project: 50-State Redistricting Simulations*. <https://doi.org/10.7910/DVN/SLCD3E>.
- Przeworski, Adam. 1974. "Contextual models of political behavior." *Political Methodology*, 27–61.

- Rivers, Douglas. 2007. "Sampling for web surveys." In *Joint Statistical Meetings*. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.468.9645>.
- Rosenman, Evan T.R., Cory McCartan, and Santiago Olivella. 2023. "Recalibration of Predicted Probabilities Using the 'Logit Shift': Why does it work, and when can it be expected to work well?" *Political Analysis*.
- Stoker, Laura, and Jake Bowers. 2002. "Designing multi-level studies: Sampling voters and electoral contexts." *Electoral Studies* 21 (2): 235–267.
- Voting and Election Science Team. 2020. *2020 Precinct-Level Election Results*. V. 35. <https://doi.org/10.7910/DVN/K7760H>.
- Yamauchi, Soichiro. 2021. *emlogit: An ECM algorithm for the Multinomial Logit Model*. <https://github.com/soichiroy/emlogit>.

Estimates from Kuriwaki et al., “Geography of Racially Polarized Voting: Calibrating Surveys at the District Level”

This section presents the estimates of two-way Republican voteshare by district by race in table form. To download the estimates in machine-readable form, refer to the Dataverse repository in <https://doi.org/10.7910/DVN/MAZNJ6>.

Congressional districts represent district lines in the 2016 and 2020 general elections. In the subsequent tables, numbers are rounded to the nearest full percentage point. Numbers enclosed by a parenthesis indicate that we estimate that the racial group comprise only 2 percent or less of the geography’s electorate. The names of the CD come from the Daily Kos **name** in August 2019. Pennsylvania redrew its Congressional districts in 2018, so we designated districts with the help of Daily Kos and Lara Putnam. The names of each district is given in the subsequent tables.

1 NATIONAL, REGION, DIVISION, AND STATE ESTIMATES

Tables **S.1** (2016) and **S.2** (2020) shows the results of these estimates at the state-level and up.

2 STANDARD ERRORS

Tables **S.3** and **S.4** show the estimated standard error of each of the state level estimates in the previous tables. They do not represent the standard errors after calibration. Tables **S.5** and **S.6** represent the standard error after calibration. See Appendix C.4 in Kuriwaki et al. (2023) for a discussion of the difference.

Table S.1 – 2016 Republican Vote by Racial Group in Regions, Divisions, and States

(a) By Region

	Non-Whites				
	White	All	Black	Hisp.	All
Northeast	51	17	4	20	43
South	69	23	8	37	55
North Central	59	19	5	29	53
West	50	26	9	25	42
National	59	22	7	29	49

(b) By Division, nested within Region

	Non-Whites				
	White	All	Black	Hisp.	All
Northeast					
New England	45	19	3	19	41
Middle Atlantic	54	16	4	21	44
South					
South Atlantic	64	21	7	37	51
East South Central	77	16	7	50	64
West South Central	72	30	9	37	58
North Central					
East North Central	57	17	5	25	51
West North Central	61	28	6	39	57
West					
Mountain	59	33	13	29	53
Pacific	44	24	9	23	36
National	59	22	7	29	49

Note: Numbers show two-party voteshare for the Republican Presidential candidate for each geography and racial group rounded to the nearest full percentage point. Columns represent, from left to right, the Republican vote among White voters, all non-White voters, Black voters, (any-part) Hispanic voters, and all racial groups in the geography. Numbers in parentheses are those where we estimate the racial group comprises only 2 percent or less of the geography's electorate. See Table S.5 for standard errors of the state-level point estimates, and Table S.11 for the estimated composition of the electorate in each state.

(c) By State, nested within Division

	Non-Whites				
	White	All	Black	Hisp.	All
New England					
Connecticut	50	20	4	24	43
Maine	49	33	(4)	(26)	49
Massachusetts	40	16	2	15	36
New Hampshire	51	32	(4)	(25)	50
Rhode Island	45	21	3	21	42
Vermont	36	21	(2)	(15)	35
Middle Atlantic					
New Jersey	55	20	4	22	43
New York	50	15	4	18	39
Pennsylvania	58	15	3	28	51
South Atlantic					
Delaware	54	14	5	26	44
Florida	63	27	6	37	51
Georgia	73	19	8	44	53
Maryland	51	13	6	25	36
North Carolina	66	19	8	39	52
South Carolina	73	17	10	50	58
Virginia	60	19	7	32	47
West Virginia	74	29	9	(54)	72
East South Central					
Alabama	83	17	7	60	65
Kentucky	71	21	3	40	66
Mississippi	84	15	10	(60)	60
Tennessee	74	15	3	44	64
West South Central					
Arkansas	74	23	8	45	65
Louisiana	81	22	13	54	61
Oklahoma	73	48	6	48	70
Texas	70	31	6	36	55
East North Central					
Illinois	51	15	5	19	41
Indiana	65	23	7	38	60
Michigan	57	17	5	31	51
Ohio	61	16	6	(34)	55
Wisconsin	54	23	3	31	51
West North Central					
Iowa	57	35	7	43	55
Kansas	65	37	8	43	61
Minnesota	53	24	4	27	50
Missouri	66	20	6	(43)	60
Nebraska	66	39	8	46	64
North Dakota	71	55	(12)	(52)	70
South Dakota	67	53	(10)	(47)	66
Mountain					
Arizona	60	28	13	23	52
Colorado	52	31	12	30	48
Idaho	69	59	(24)	53	69
Montana	62	54	(16)	(38)	62
Nevada	57	32	12	28	49
New Mexico	55	34	(14)	32	46
Utah	65	48	(18)	39	63
Wyoming	77	63	(27)	59	76
Pacific					
Alaska	62	47	16	45	59
California	43	23	9	22	34
Hawaii	42	26	9	23	33
Oregon	46	29	(6)	27	44
Washington	45	28	7	30	42
National	59	22	7	29	49

Table S.2 – 2020 Republican Vote by Race and Geography

(a) Region Level						(c) State Level					
	White	Non-Whites			All		White	Non-Whites			All
		All	Black	Hisp.				All	Black	Hisp.	
Northeast	48	22	10	29	42						
South	66	27	10	44	53						
North Central	56	24	9	37	51						
West	48	31	12	34	42						
National	56	28	10	38	48						
(b) Division Level											
	White	Non-Whites			All		White	Non-Whites			All
		All	Black	Hisp.				All	Black	Hisp.	
Northeast											
New England	40	25	12	28	38						
Middle Atlantic	52	22	9	29	44						
South											
South Atlantic	62	25	10	43	49						
East South Central	73	22	11	60	62						
West South Central	69	34	11	44	56						
North Central											
East North Central	56	21	8	35	50						
West North Central	58	32	14	42	55						
West											
Mountain	56	39	16	40	51						
Pacific	43	29	10	32	37						
National	56	28	10	38	48						
						New England					
						Connecticut	44	28	15	34	40
						Maine	46	33	(12)	(38)	46
						Massachusetts	36	22	10	22	33
						New Hampshire	47	36	(16)	(33)	46
						Rhode Island	42	28	11	33	40
						Vermont	32	22	(8)	(24)	32
						Middle Atlantic					
						New Jersey	50	27	11	34	42
						New York	49	20	9	26	40
						Pennsylvania	56	19	8	32	50
						South Atlantic					
						Delaware	48	22	12	33	41
						Florida	62	34	10	46	52
						Georgia	70	20	9	45	50
						Maryland	46	15	6	25	33
						North Carolina	62	24	13	40	51
						South Carolina	70	20	10	59	56
						Virginia	56	22	9	33	45
						West Virginia	72	38	16	(51)	70
						East South Central					
						Alabama	79	22	12	71	63
						Kentucky	68	26	11	(48)	63
						Mississippi	82	19	13	(73)	59
						Tennessee	70	22	7	54	62
						West South Central					
						Arkansas	72	30	12	55	64
						Louisiana	79	23	13	61	60
						Oklahoma	71	47	14	62	67
						Texas	65	35	9	42	53
						East North Central					
						Illinois	50	19	6	32	42
						Indiana	62	28	14	42	58
						Michigan	55	21	7	40	49
						Ohio	59	19	9	(38)	54
						Wisconsin	53	27	8	35	50
						West North Central					
						Iowa	56	37	16	46	54
						Kansas	61	41	20	51	58
						Minnesota	50	25	11	30	46
						Missouri	63	26	12	(44)	58
						Nebraska	63	41	20	49	60
						North Dakota	69	51	(27)	(53)	67
						South Dakota	64	55	(22)	(57)	64
						Mountain					
						Arizona	57	36	15	35	50
						Colorado	47	33	15	35	43
						Idaho	67	62	(29)	59	66
						Montana	59	54	(30)	(53)	59
						Nevada	55	41	16	50	49
						New Mexico	51	39	(19)	40	45
						Utah	63	51	(27)	48	61
						Wyoming	73	66	(34)	67	73
						Pacific					
						Alaska	58	48	17	46	55
						California	42	28	10	31	35
						Hawaii	42	31	14	26	35
						Oregon	44	32	(9)	33	42
						Washington	43	31	10	34	40
						National	56	28	10	38	48

Note: See Table S.1

Table S.3 – Standard Errors of 2016 Estimates, Hierarchical Model.

		Non-Whites				All
		White	All	Black	Hisp.	
New England						
Connecticut	.021	.031	.024	.064	.019	
Maine	.023	.047	.024	.076	.023	
Massachusetts	.015	.023	.016	.041	.014	
New Hampshire	.023	.043	.026	.075	.023	
Rhode Island	.027	.041	.024	.076	.025	
Vermont	.03	.044	.019	.068	.03	
Middle Atlantic						
New Jersey	.016	.021	.016	.042	.013	
New York	.012	.014	.013	.028	.0099	
Pennsylvania	.011	.014	.013	.046	.01	
South Atlantic						
Delaware	.026	.024	.022	.08	.022	
Florida	.01	.016	.014	.03	.0092	
Georgia	.015	.015	.015	.059	.012	
Maryland	.017	.015	.016	.054	.013	
North Carolina	.015	.015	.015	.053	.012	
South Carolina	.019	.019	.021	.071	.016	
Virginia	.014	.017	.015	.05	.012	
West Virginia	.021	.026	.029	.082	.02	
East South Central						
Alabama	.02	.017	.018	.078	.016	
Kentucky	.02	.021	.012	.076	.018	
Mississippi	.023	.022	.024	.085	.017	
Tennessee	.017	.015	.014	.068	.015	
West South Central						
Arkansas	.021	.024	.023	.08	.019	
Louisiana	.019	.024	.028	.069	.016	
Oklahoma	.02	.03	.023	.077	.018	
Texas	.012	.019	.014	.029	.011	
East North Central						
Illinois	.013	.015	.014	.034	.011	
Indiana	.016	.021	.018	.059	.015	
Michigan	.014	.016	.014	.05	.013	
Ohio	.012	.014	.015	.051	.011	
Wisconsin	.016	.023	.015	.057	.015	
West North Central						
Iowa	.019	.04	.031	.08	.018	
Kansas	.021	.038	.03	.079	.02	
Minnesota	.017	.028	.021	.058	.016	
Missouri	.016	.02	.022	.064	.014	
Nebraska	.022	.041	.032	.083	.021	
North Dakota	.03	.05	.048	.12	.03	
South Dakota	.03	.058	.052	.11	.029	
Mountain						
Arizona	.017	.029	.043	.04	.015	
Colorado	.017	.031	.041	.044	.016	
Idaho	.025	.054	.081	.086	.025	
Montana	.029	.056	.074	.11	.028	
Nevada	.021	.033	.046	.055	.02	
New Mexico	.024	.042	.062	.055	.025	
Utah	.023	.044	.063	.065	.022	
Wyoming	.031	.064	.087	.1	.031	
Pacific						
Alaska	.03	.057	.07	.11	.03	
California	.011	.015	.026	.023	.0095	
Hawaii	.03	.052	.051	.077	.033	
Oregon	.018	.033	.027	.057	.017	
Washington	.015	.025	.026	.05	.014	

Table S.4 – Standard Errors of 2020 Estimates, Hierarchical Model.

		Non-Whites			
	White	All	Black	Hisp.	All
New England					
Connecticut	.02	.034	.049	.064	.019
Maine	.024	.044	.054	.084	.024
Massachusetts	.015	.027	.036	.045	.015
New Hampshire	.023	.04	.063	.064	.023
Rhode Island	.026	.041	.045	.075	.025
Vermont	.026	.043	.044	.069	.026
Middle Atlantic					
New Jersey	.015	.02	.02	.04	.013
New York	.011	.014	.013	.028	.0095
Pennsylvania	.01	.014	.015	.041	.0093
South Atlantic					
Delaware	.03	.035	.044	.082	.026
Florida	.011	.014	.013	.026	.0089
Georgia	.016	.012	.01	.051	.011
Maryland	.018	.013	.012	.045	.013
North Carolina	.014	.014	.015	.046	.012
South Carolina	.019	.015	.016	.066	.015
Virginia	.014	.015	.013	.043	.012
West Virginia	.023	.031	.042	.083	.022
East South Central					
Alabama	.02	.016	.018	.074	.016
Kentucky	.019	.025	.029	.075	.017
Mississippi	.028	.021	.023	.08	.021
Tennessee	.016	.016	.015	.071	.014
West South Central					
Arkansas	.023	.024	.024	.081	.021
Louisiana	.024	.018	.019	.067	.018
Oklahoma	.023	.03	.034	.076	.021
Texas	.011	.015	.012	.023	.0095
East North Central					
Illinois	.013	.016	.011	.039	.011
Indiana	.015	.021	.024	.053	.014
Michigan	.013	.016	.014	.049	.012
Ohio	.012	.013	.014	.046	.011
Wisconsin	.016	.022	.021	.053	.015
West North Central					
Iowa	.02	.037	.041	.075	.019
Kansas	.023	.035	.045	.07	.022
Minnesota	.016	.025	.035	.051	.015
Missouri	.015	.022	.026	.057	.014
Nebraska	.025	.038	.051	.074	.024
North Dakota	.031	.051	.084	.097	.031
South Dakota	.031	.052	.078	.092	.031
Mountain					
Arizona	.017	.027	.037	.041	.016
Colorado	.016	.025	.035	.04	.015
Idaho	.026	.048	.076	.073	.025
Montana	.027	.057	.088	.09	.027
Nevada	.021	.032	.044	.061	.02
New Mexico	.026	.037	.067	.05	.026
Utah	.025	.043	.072	.065	.024
Wyoming	.034	.058	.093	.093	.033
Pacific					
Alaska	.035	.055	.057	.086	.034
California	.01	.013	.017	.021	.009
Hawaii	.032	.043	.057	.061	.032
Oregon	.016	.028	.03	.048	.016
Washington	.015	.023	.025	.046	.014

Table S.5 – Standard Errors of 2016 Estimates, after Calibration. Standard errors after calibration.

		Non-Whites		
	White	All	Black	Hisp.
New England				
Connecticut	.0069	.024	.024	.053
Maine	.0015	.043	.025	.07
Massachusetts	.0043	.019	.016	.033
New Hampshire	.0021	.04	.029	.071
Rhode Island	.0065	.033	.023	.063
Vermont	.0043	.039	.019	.057
Middle Atlantic				
New Jersey	.0077	.015	.016	.033
New York	.0053	.01	.012	.021
Pennsylvania	.0026	.012	.012	.039
South Atlantic				
Delaware	.0067	.018	.019	.068
Florida	.0056	.011	.013	.023
Georgia	.0072	.012	.015	.054
Maryland	.0072	.011	.013	.044
North Carolina	.0053	.013	.016	.049
South Carolina	.0066	.017	.02	.07
Virginia	.0056	.013	.014	.043
West Virginia	.0027	.027	.034	.081
East South Central				
Alabama	.008	.019	.023	.072
Kentucky	.0025	.02	.014	.074
Mississippi	.012	.026	.03	.082
Tennessee	.0029	.015	.015	.065
West South Central				
Arkansas	.005	.022	.025	.076
Louisiana	.012	.021	.028	.066
Oklahoma	.0044	.025	.026	.074
Texas	.007	.011	.015	.019
East North Central				
Illinois	.0045	.012	.013	.028
Indiana	.0027	.02	.021	.058
Michigan	.0029	.015	.015	.047
Ohio	.0023	.015	.016	.05
Wisconsin	.0026	.021	.016	.053
West North Central				
Iowa	.0031	.037	.035	.078
Kansas	.0047	.034	.036	.076
Minnesota	.0035	.025	.022	.053
Missouri	.0033	.022	.027	.064
Nebraska	.0044	.038	.041	.08
North Dakota	.0045	.042	.063	.12
South Dakota	.0043	.05	.061	.11
Mountain				
Arizona	.0073	.021	.043	.031
Colorado	.0069	.023	.042	.036
Idaho	.0048	.049	.094	.083
Montana	.0042	.05	.081	.1
Nevada	.011	.023	.046	.044
New Mexico	.019	.023	.058	.037
Utah	.005	.038	.073	.061
Wyoming	.012	.058	.11	.1
Pacific				
Alaska	.013	.041	.072	.1
California	.0064	.0072	.021	.013
Hawaii	.031	.022	.037	.057
Oregon	.004	.027	.025	.049
Washington	.004	.02	.024	.043

Table S.6 – Standard Errors of 2020 Estimates, after Calibration.

		Non-Whites		
	White	All	Black	Hisp.
New England				
Connecticut	.0094	.03	.062	.059
Maine	.0027	.041	.07	.082
Massachusetts	.0053	.022	.045	.042
New Hampshire	.0033	.038	.085	.066
Rhode Island	.0073	.035	.061	.07
Vermont	.0083	.039	.061	.069
Middle Atlantic				
New Jersey	.0074	.015	.026	.036
New York	.0057	.011	.019	.027
Pennsylvania	.0032	.015	.021	.041
South Atlantic				
Delaware	.014	.033	.053	.076
Florida	.006	.011	.019	.021
Georgia	.0077	.011	.015	.05
Maryland	.0076	.011	.014	.042
North Carolina	.0063	.014	.021	.047
South Carolina	.0068	.018	.023	.062
Virginia	.006	.013	.018	.042
West Virginia	.0033	.037	.062	.083
East South Central				
Alabama	.01	.023	.03	.063
Kentucky	.0038	.032	.047	.077
Mississippi	.015	.029	.034	.068
Tennessee	.0041	.02	.025	.07
West South Central				
Arkansas	.0067	.027	.042	.076
Louisiana	.013	.021	.03	.062
Oklahoma	.0049	.026	.049	.07
Texas	.0071	.01	.019	.018
East North Central				
Illinois	.005	.014	.015	.036
Indiana	.0034	.024	.038	.054
Michigan	.0035	.016	.02	.049
Ohio	.0026	.017	.023	.05
Wisconsin	.0031	.022	.032	.054
West North Central				
Iowa	.0031	.035	.064	.075
Kansas	.006	.032	.067	.067
Minnesota	.0043	.025	.049	.052
Missouri	.0041	.027	.04	.059
Nebraska	.0065	.036	.079	.073
North Dakota	.0073	.049	.12	.098
South Dakota	.005	.045	.11	.087
Mountain				
Arizona	.0089	.018	.047	.033
Colorado	.0074	.02	.05	.036
Idaho	.0062	.037	.11	.066
Montana	.0056	.049	.12	.088
Nevada	.014	.02	.056	.051
New Mexico	.019	.018	.079	.035
Utah	.0076	.033	.099	.061
Wyoming	.014	.048	.13	.082
Pacific				
Alaska	.014	.038	.078	.083
California	.0067	.0065	.022	.013
Hawaii	.029	.017	.071	.057
Oregon	.0048	.025	.043	.046
Washington	.0051	.019	.036	.044

3 CONGRESSIONAL DISTRICT ESTIMATES

Tables S.7 and S.8 shows the estimated Trump vote within each in each congressional district. For example, the first number in the table means that 62 percent of all (election participating, two-party) Whites in Alaska’s at-large congressional district are estimated to have voted for Trump over Clinton in 2016.

The CD-level estimates are calibrated to the actual election result so can be interpreted as the actual two-party Trump vote in that CD. Among non-Whites, we provide an “All” (non-Whites) aggregate as well.

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD

	White	Non-Whites			
		All	Black	Hisp.	All
All of Alaska: AK-01	62	47	16	45	59
Greater Mobile: AL-01	83	18	7	(64)	65
Southeastern Alabama: AL-02	87	16	9	(67)	67
Eastern Alabama: AL-03	85	16	8	(65)	67
North-central Alabama: AL-04	87	41	8	69	83
Northern Alabama: AL-05	78	22	5	53	68
Birmingham suburbs: AL-06	84	28	7	60	73
Birmingham and the Black Belt: AL-07	66	8	5	(40)	29
Northeastern Arkansas: AR-01	77	19	9	(55)	69
Greater Little Rock: AR-02	68	15	6	42	56
Northwest Arkansas: AR-03	71	44	6	43	67
Southern and western Arkansas: AR-04	79	21	9	52	68
Northeastern Arizona: AZ-01	59	33	(15)	24	51
Southeastern Arizona and eastern Tucson area: AZ-02	55	26	13	21	48
Southwestern Arizona and western Tucson area: AZ-03	50	20	12	15	35
North-central Arizona and Phoenix exurbs: AZ-04	75	47	(24)	44	71
Phoenix suburbs of Mesa and Gilbert: AZ-05	66	38	(19)	33	62
Scottsdale and North Phoenix: AZ-06	59	36	(15)	34	56
Downtown and West Phoenix: AZ-07	36	14	7	10	24
Western suburbs of Phoenix: AZ-08	67	36	18	33	61
Central Phoenix and eastern suburbs: AZ-09	48	22	10	17	42
Northeastern California: CA-01	63	44	(18)	41	61
Northern California coast: CA-02	27	19	(5)	17	26
Southern Sacramento Valley: CA-03	51	31	13	31	44
Northern Sacramento suburbs and Sierra Nevada foothi...: CA-04	60	45	(17)	41	58
Wine country: CA-05	31	18	6	17	26
Sacramento: CA-06	34	17	7	17	26
Suburban Sacramento County: CA-07	49	34	12	40	44
Northern San Bernardino County and the High Desert: CA-08	68	41	23	40	58
Stockton area and eastern Contra Costa County: CA-09	51	29	12	25	41

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD (*continued*)

	White	Non-Whites			All
		All	Black	Hisp.	
Modesto area: CA-10	57	35	14	32	49
Contra Costa County: CA-11	30	17	6	17	24
San Francisco: CA-12	12	6	2	7	9
Oakland area: CA-13	11	5	3	5	8
San Mateo County: CA-14	26	14	4	15	19
Eastern Alameda County: CA-15	34	19	7	20	26
Fresno area: CA-16	55	28	14	23	39
Northeast San Jose area: CA-17	31	16	6	18	22
Silicon Valley: CA-18	25	16	(5)	16	22
San Jose: CA-19	32	18	6	16	23
Monterey Bay: CA-20	32	18	(6)	14	25
Southern Central Valley and part of Bakersfield: CA-21	59	32	17	28	42
Central Valley near Fresno and Tulare: CA-22	65	42	20	37	55
Bakersfield area: CA-23	70	47	26	45	62
San Luis Obispo and Santa Barbara: CA-24	45	27	(10)	25	40
Northern Los Angeles County suburbs: CA-25	57	35	16	32	47
Ventura County: CA-26	46	27	(11)	24	39
San Gabriel Valley: CA-27	42	23	10	24	30
Burbank and Glendale: CA-28	27	17	5	16	24
East San Fernando Valley: CA-29	28	13	7	11	18
West San Fernando Valley: CA-30	32	19	6	17	27
San Bernardino area: CA-31	54	29	15	27	39
East San Gabriel Valley: CA-32	46	24	10	19	30
Coastal West Side of Los Angeles: CA-33	31	22	5	26	28
Downtown Los Angeles: CA-34	22	10	4	8	12
Southwest San Bernardino County and Pomona: CA-35	47	23	12	19	29
Eastern Riverside County: CA-36	55	30	14	25	46
Culver City area of Los Angeles: CA-37	16	8	4	9	11
Southeastern Los Angeles County: CA-38	46	24	10	21	29
San Gabriel Valley and northern Orange County: CA-39	58	36	16	32	46
East Los Angeles: CA-40	31	12	5	10	14
Riverside and Moreno Valley: CA-41	52	27	13	23	36
Southwestern Riverside County: CA-42	66	44	22	40	57
Inland cities south of Los Angeles: CA-43	31	14	8	13	18
South Los Angeles, including Compton: CA-44	31	11	6	9	13
Irvine area of Orange County: CA-45	53	36	(14)	32	47
Santa Ana and the Anaheim area of Orange County: CA-46	43	23	9	20	30
Long Beach area: CA-47	43	24	10	24	33
Coastal Orange County: CA-48	53	41	(14)	42	49
Southern Orange and northern San Diego Counties: CA-49	50	35	(12)	34	46
Inland San Diego County: CA-50	63	46	19	45	58
South San Diego and Imperial County: CA-51	39	20	8	19	24
Suburban San Diego: CA-52	42	28	10	24	38
Eastern San Diego: CA-53	39	23	9	23	32

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Denver: CO-01	31	15	6	15	25
Boulder and Fort Collins: CO-02	40	25	(9)	22	39
Pueblo and the Western Slope: CO-03	62	40	(15)	36	57
Eastern Colorado and Denver exurbs: CO-04	67	46	(19)	43	63
Colorado Springs: CO-05	68	46	24	43	64
Aurora and Denver's southern suburbs: CO-06	51	32	12	37	45
Denver's northern and western suburbs: CO-07	49	28	(10)	25	44
Hartford area: CT-01	47	17	4	23	38
Eastern Connecticut: CT-02	51	27	4	26	49
New Haven area: CT-03	50	18	4	24	42
Southern Fairfield County: CT-04	47	18	4	21	38
Northern Fairfield County and northwestern Connecticut: CT-05	53	28	4	30	48
All of Delaware: DE-01	54	14	5	26	44
Pensacola area: FL-01	79	38	13	62	71
Florida panhandle: FL-02	77	30	11	59	69
North Florida, including Gainesville and Ocala: FL-03	68	28	8	53	59
Part of the Jacksonville area and St. Augustine: FL-04	71	38	8	57	65
Jacksonville and Tallahassee areas: FL-05	61	15	9	41	37
Daytona area: FL-06	65	32	6	47	59
Northern Orlando suburbs: FL-07	54	30	4	38	47
Space Coast: FL-08	67	32	8	46	61
Southern suburbs of Orlando: FL-09	60	26	5	27	44
Orlando: FL-10	53	22	5	32	36
North-central Florida and The Villages: FL-11	71	37	7	51	67
Northern Tampa suburbs: FL-12	64	35	6	39	60
St. Petersburg area: FL-13	54	21	5	32	49
Tampa area: FL-14	53	25	5	33	41
Lakeland and the exurbs of Tampa and Orlando: FL-15	66	32	7	43	56
Sarasota area: FL-16	60	33	5	45	56
South-central Florida: FL-17	69	35	7	47	64
Treasure Coast and Palm Beach area: FL-18	62	28	6	42	55
Cape Coral and Fort Myers: FL-19	66	36	6	45	62
West Palm Beach and Fort Lauderdale areas: FL-20	39	11	4	24	19
Southeast Palm Beach County: FL-21	47	22	3	33	40
Boca Raton and northeast Broward County: FL-22	49	24	3	36	42
Broward County-centered: FL-23	46	27	3	35	37
Northern Miami: FL-24	36	12	4	20	16
Northwestern Miami suburbs and eastern Collier Count...: FL-25	75	41	10	42	51
Southwestern Miami area and the Florida Keys: FL-26	61	36	6	40	42
Southern Miami area and Coral Gables: FL-27	59	33	6	34	40
Coastal Georgia: GA-01	75	19	10	54	58
Southwestern Georgia: GA-02	73	15	11	47	44

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD (*continued*)

	White	Non-Whites			All
		All	Black	Hisp.	
Southwestern Atlanta exurbs: GA-03	80	26	14	59	67
Eastern Atlanta suburbs: GA-04	51	10	5	27	23
Downtown Atlanta: GA-05	28	4	2	14	13
Northern Atlanta suburbs: GA-06	62	29	6	37	51
Northeastern Atlanta suburbs: GA-07	71	32	8	41	54
South-central Georgia: GA-08	83	24	15	65	65
Northeastern Georgia: GA-09	85	45	14	64	81
East-Central Georgia: GA-10	78	23	12	58	63
Northwestern Atlanta suburbs and exurbs: GA-11	75	31	8	47	63
Augusta and southeastern Georgia: GA-12	78	21	14	56	59
Southwestern Atlanta suburbs: GA-13	59	12	6	34	28
Northwestern Georgia: GA-14	84	38	12	63	78
Honolulu: HI-01	46	26	10	28	33
Northern Oahu and all other islands: HI-02	40	25	(6)	19	33
Northeastern Iowa: IA-01	54	30	6	(38)	52
Southeastern Iowa: IA-02	55	31	6	39	53
Des Moines and southwestern Iowa: IA-03	54	33	6	43	52
Northwestern Iowa: IA-04	66	47	(9)	50	65
Idaho Panhandle and western Boise area: ID-01	72	63	(27)	56	72
Boise and Eastern Idaho: ID-02	66	56	(22)	52	65
South Side Chicago and southern suburbs: IL-01	42	9	6	22	22
South Side Chicago and southwestern suburbs: IL-02	40	8	5	19	20
Southwestern Chicago area: IL-03	50	23	4	21	42
Heavily Latino areas of Chicago's West and South sid...: IL-04	24	9	1	7	14
North Side Chicago: IL-05	29	14	2	12	26
Western Chicago suburbs: IL-06	50	30	(4)	28	47
West Side and Downtown Chicago: IL-07	19	4	2	9	10
Northwestern Chicago suburbs: IL-08	45	23	3	21	39
Northern Chicago Suburbs: IL-09	30	14	2	15	27
Chicago's North Shore suburbs: IL-10	40	19	3	19	35
Southwestern Chicago suburbs: IL-11	47	20	4	22	38
Southwestern Illinois: IL-12	65	17	9	(40)	58
Central Illinois: IL-13	58	19	6	(32)	53
Western Chicago exurbs: IL-14	56	31	5	28	52
East-central and southeastern Illinois: IL-15	77	32	12	(54)	75
North-central Illinois: IL-16	62	31	6	37	60
Northwestern Illinois: IL-17	55	18	5	32	51
West-central Illinois: IL-18	67	33	8	(41)	65
Northwestern Indiana: IN-01	52	18	6	27	44
North-central Indiana: IN-02	66	27	7	42	63
Northeastern Indiana: IN-03	72	33	10	53	69
West-central Indiana: IN-04	71	39	9	53	68

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD (*continued*)

		Non-Whites				All
		White	All	Black	Hisp.	
South-central Indiana and southern Indianapolis suburbs: IN-09	Northern Indianapolis suburbs: IN-05	61	22	6	41	57
	Southeastern Indiana: IN-06	73	35	(10)	(53)	72
	Indianapolis: IN-07	47	13	5	29	38
	Southwestern Indiana: IN-08	70	29	9	(52)	68
	Western Kansas: KS-01	66	34	8	(44)	64
	Topeka, Lawrence, and eastern rural Kansas: KS-02	76	53	(14)	55	74
	Suburban Kansas City area: KS-03	62	36	8	37	60
	Wichita area: KS-04	54	28	5	37	50
	Western Kentucky: KY-01	68	42	10	47	65
Cincinnati suburbs and northern Kentucky border area: KY-04	West-central Kentucky: KY-02	80	26	5	(57)	76
	Louisville: KY-03	75	31	4	(51)	71
	Appalachian Kentucky: KY-05	53	11	2	27	42
	Greater Lexington area: KY-06	72	32	4	(43)	69
		83	41	(7)	(67)	82
		63	23	2	38	58
New Orleans and Baton Rouge: LA-02	Suburban New Orleans: LA-01	81	39	11	58	72
	Southwestern Louisiana: LA-03	50	10	6	24	23
	Northwestern Louisiana: LA-04	85	28	16	65	70
	Northeastern Louisiana: LA-05	83	24	16	63	63
	Baton Rouge suburbs: LA-06	87	25	19	(69)	66
		82	28	14	58	68
Boston's northern and outer western suburbs: MA-05	Springfield area and the Berkshires: MA-01	44	17	4	16	39
	Worcester area and upper Pioneer Valley: MA-02	42	21	3	18	40
	Lowell area and the Merrimack Valley: MA-03	43	22	(3)	17	38
	North Shore: MA-06	39	22	(3)	19	38
	Central Boston area: MA-07	31	13	2	11	27
	Southern Boston suburbs: MA-08	44	21	(3)	17	41
	Cape Cod and South Shore: MA-09	18	6	1	6	13
		41	17	3	20	37
		46	25	(3)	(19)	45
Baltimore suburbs to Annapolis to D.C. suburbs: MD-03	Eastern Shore and Baltimore exurbs: MD-01	71	22	9	(45)	65
	Suburban Baltimore area: MD-02	51	12	5	29	38
	Prince George's County D.C. suburbs: MD-04	43	13	4	22	34
	Southern Maryland and outer Prince George's County: MD-05	43	11	6	20	21
	Western Maryland and Montgomery County D.C. suburbs: MD-06	53	14	7	29	34
	Baltimore and western suburbs: MD-07	50	23	4	29	42
	Montgomery County: MD-08	40	8	6	20	21
		41	16	3	21	32
Upper Peninsula and northern Lower Peninsula: MI-01	Portland and Augusta: ME-01	43	27	(3)	(22)	42
	Northern Maine: ME-02	56	40	(5)	(31)	56

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD (*continued*)

		Non-Whites			
		White	All	Black	Hisp. All
	Western Michigan coast: MI-02	64	32	7	38 60
	Grand Rapids area: MI-03	59	25	6	36 55
	Central Michigan: MI-04	65	35	(7)	(36) 63
	Flint, Saginaw, and Bay City: MI-05	55	13	5	32 48
	Southwestern Michigan: MI-06	58	23	6	34 55
	South-central Michigan: MI-07	61	28	6	(38) 59
	Lansing and northwestern Detroit exurbs: MI-08	57	27	5	30 54
	Northern Detroit suburbs: MI-09	51	16	4	(28) 46
	Northern Macomb County and the "Thumb" region: MI-10	69	37	(8)	(43) 67
	Northwestern Detroit suburbs: MI-11	56	31	5	(34) 53
	Ann Arbor and Dearborn: MI-12	40	16	3	20 37
	Western Detroit: MI-13	33	6	4	15 19
	Eastern Detroit to Pontiac: MI-14	38	7	5	18 19
	Southern Minnesota: MN-01	60	38	(9)	37 58
	Southern Twin Cities suburbs: MN-02	53	33	6	38 51
	Western Twin Cities suburbs: MN-03	48	25	4	25 45
	St. Paul: MN-04	38	19	4	20 34
	Minneapolis: MN-05	24	9	2	11 20
	Northern Twin Cities exurbs: MN-06	65	41	(12)	(43) 64
	Western Minnesota: MN-07	68	49	(10)	(45) 67
	Iron Range: MN-08	59	43	(8)	(35) 59
	St. Louis and northwestern suburbs: MO-01	32	6	4	19 20
	Suburban St. Louis: MO-02	58	34	7	(41) 56
	East-central Missouri: MO-03	72	37	11	(55) 71
	West-central Missouri: MO-04	72	36	10	(54) 69
	Kansas City area: MO-05	51	16	6	31 43
	Northern Missouri: MO-06	69	35	10	(52) 67
	Southwestern Missouri: MO-07	76	51	(13)	(57) 74
	Southeastern Missouri: MO-08	81	37	16	(61) 79
	Northeastern Mississippi: MS-01	85	16	10	(61) 67
	Mississippi Delta and Jackson: MS-02	77	12	10	(53) 36
	Southwestern and eastern Mississippi: MS-03	85	17	11	(63) 63
	Gulf Coast: MS-04	86	20	9	(62) 72
	All of Montana: MT-01	62	54	(16)	(38) 62
	Durham and northeastern North Carolina: NC-01	53	11	7	31 31
	Raleigh exurbs: NC-02	68	24	10	45 55
	Eastern North Carolina coast: NC-03	75	21	12	51 62
	Raleigh, Cary, Chapel Hill: NC-04	41	11	3	21 30
	Winston-Salem and northwestern North Carolina: NC-05	68	22	9	42 59
	Part of Greensboro and rural Piedmont: NC-06	71	22	9	47 58
	Southeastern North Carolina: NC-07	71	21	10	51 59
	Charlotte suburbs to Fayetteville: NC-08	72	24	12	48 58

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
South Charlotte to suburban Fayetteville: NC-09	69	28	9	48	56
Asheville and southwestern Piedmont: NC-10	70	25	11	47	63
Appalachian North Carolina: NC-11	68	35	7	47	65
Charlotte: NC-12	48	12	5	26	30
Greensboro and rural western Piedmont: NC-13	69	20	9	43	55
All of North Dakota: ND-01	71	55	(12)	(52)	70
Lincoln and rural eastern Nebraska: NE-01	63	43	(8)	44	62
Greater Omaha area: NE-02	56	27	7	36	52
Western Nebraska: NE-03	80	63	(16)	61	79
Eastern New Hampshire: NH-01	52	32	(5)	(25)	51
Western and northern New Hampshire: NH-02	50	32	(4)	(25)	49
Inner Philadelphia suburbs: NJ-01	47	16	4	22	38
Southern New Jersey coast: NJ-02	62	25	7	31	53
Philadelphia suburbs and central Jersey Shore: NJ-03	60	23	6	31	54
Monmouth County area: NJ-04	63	29	7	35	58
Northern Bergen County and northwestern exurbs: NJ-05	58	30	5	29	51
Northern Middlesex County and northern Jersey Shore: NJ-06	53	24	5	26	42
Hunterdon County and New York City's western exurbs: NJ-07	56	28	5	28	50
Jersey City and inner New York City suburbs: NJ-08	36	15	2	15	22
Paterson and inner New York City suburbs: NJ-09	49	21	4	19	34
Newark area: NJ-10	30	7	3	14	13
Morris County area: NJ-11	56	29	4	28	51
Trenton and southern Middlesex County: NJ-12	45	16	4	19	33
Albuquerque area: NM-01	50	30	(11)	28	41
Southern New Mexico: NM-02	65	44	(18)	42	56
Northern New Mexico: NM-03	52	32	(14)	28	42
Las Vegas: NV-01	46	21	8	17	35
Northern Nevada: NV-02	62	40	(14)	33	57
Southern Las Vegas suburbs: NV-03	56	37	12	36	51
Northern Las Vegas suburbs and rural central Nevada: NV-04	58	30	14	30	48
Eastern Suffolk County: NY-01	62	30	6	32	57
Southern Long Island: NY-02	65	31	8	33	55
Northern Long Island: NY-03	52	32	4	30	47
Long Island: NY-04	57	28	6	33	45
Southeastern Queens: NY-05	34	10	5	15	13
Northeastern Queens: NY-06	42	23	3	24	33
Heavily Latino parts of northern Brooklyn and southw...: NY-07	18	7	1	7	11
Eastern Brooklyn: NY-08	29	7	3	11	14
Central Brooklyn: NY-09	28	7	3	14	15
West Side of Manhattan and part of southern Brooklyn: NY-10	23	10	1	9	20
Staten Island and part of southern Brooklyn: NY-11	62	35	6	41	55

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD (*continued*)

		Non-Whites			
		White	All	Black	Hisp. All
East Side of Manhattan and nearby parts of Queens an...:	NY-12	16	8	1	9 14
Harlem area:	NY-13	16	5	1	7 7
East Bronx and northern Queens:	NY-14	34	13	2	12 20
The Bronx:	NY-15	24	7	2	6 7
Southern Westchester County and the northern Bronx:	NY-16	39	11	3	16 23
Westchester County:	NY-17	50	22	4	21 40
Lower Hudson Valley:	NY-18	59	27	6	33 51
Hudson Valley:	NY-19	57	26	5	32 54
Albany area:	NY-20	47	18	4	30 43
North Country:	NY-21	59	27	6	(35) 58
Binghamton and Utica-Rome:	NY-22	61	25	6	(33) 59
Ithaca and southwestern New York:	NY-23	60	27	5	(33) 58
Syracuse area:	NY-24	52	16	4	(25) 48
Greater Rochester:	NY-25	49	16	5	27 42
Greater Buffalo:	NY-26	48	12	4	25 40
Suburbs and rural areas between Buffalo and Rochester:	NY-27	65	30	(7)	(40) 63
Suburban Cincinnati:	OH-01	64	18	10	(41) 54
Eastern Cincinnati and east along the Ohio River:	OH-02	62	19	9	(41) 59
Columbus:	OH-03	41	9	4	24 30
West-Central Ohio to North-Central Ohio:	OH-04	71	29	10	(48) 68
Northwestern Ohio:	OH-05	66	35	(9)	42 64
Appalachian Ohio:	OH-06	74	32	(11)	(56) 73
Canton area:	OH-07	68	27	9	(47) 66
Cincinnati suburbs and Springfield:	OH-08	72	30	10	(50) 68
Toledo to West Side of Cleveland:	OH-09	45	13	4	23 39
Greater Dayton area:	OH-10	61	16	8	(39) 54
East Cleveland to Akron:	OH-11	31	6	4	16 18
Northern Columbus area:	OH-12	59	27	7	(34) 56
Mahoning Valley and Akron:	OH-13	51	13	5	(30) 47
Northeastern Ohio:	OH-14	59	23	6	(35) 56
Southern Columbus area and Athens:	OH-15	60	26	7	(39) 58
Western suburbs of Cleveland and Akron:	OH-16	60	32	(6)	(42) 59
Greater Tulsa area:	OK-01	70	45	6	48 66
Eastern Oklahoma:	OK-02	79	62	8	55 77
Northwestern Oklahoma:	OK-03	81	59	9	60 78
Southern Oklahoma:	OK-04	73	48	7	49 70
Oklahoma City area:	OK-05	63	34	5	37 58
Western Portland suburbs and North Coast:	OR-01	40	28	(7)	25 38
Oregon east of the Cascades and part of southern Oregon:	OR-02	63	45	(14)	40 61
Eastern Portland area:	OR-03	27	14	4	12 24
Southern Willamette Valley and southern coast:	OR-04	51	36	(10)	29 50
Salem area and southern Portland suburbs:	OR-05	50	34	(10)	31 48
Philadelphia along the Delaware River and City of Ch...:	PA-01	31	7	2	13 19

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

		Non-Whites				All
		White	All	Black	Hisp.	
Philadelphia City center and Chestnut Hill sububrs:	PA-02	20	3	2	9	9
Erie and northwestern Pennsylvania:	PA-03	66	23	6	(45)	64
Harrisburg and York:	PA-04	66	26	7	40	62
North-central Pennsylvania through Centre County:	PA-05	67	29	7	(42)	65
Philadelphia’s Chester County suburbs and Reading su....:	PA-06	53	24	4	29	50
Philadelphia’s Delaware and Montgomery County suburbs:	PA-07	53	21	4	(32)	49
Philadelphia’s Bucks County suburbs:	PA-08	53	26	4	32	50
South-central Pennsylvania:	PA-09	74	30	10	(47)	72
Williamsport area and northeastern Pennsylvania:	PA-10	71	30	8	(46)	69
Coal Region between Wilkes-Barre and Harrisburg:	PA-11	66	27	7	39	63
Beaver County through Pittsburgh’s North Hills to Jo....:	PA-12	63	23	6	(36)	61
Northeast Philadelphia and Norristown:	PA-13	42	14	3	21	33
Pittsburgh:	PA-14	39	7	3	(21)	32
Lehigh Valley stretched to Susquehanna River:	PA-15	59	29	5	31	54
Reading, Lancaster, and Coatesville:	PA-16	59	29	5	36	54
Scranton/Wilkes-Barre and the Slate Belt:	PA-17	59	23	6	32	56
Pittsburgh’s South and West Hills, Washington, Westm....:	PA-18	62	26	(6)	(37)	61
Part of Providence and eastern Rhode Island:	RI-01	41	18	3	19	37
Part of Providence and western Rhode Island:	RI-02	49	25	4	23	47
Charleston and coastal South Carolina:	SC-01	68	17	8	47	57
Augusta and Columbia suburbs:	SC-02	73	19	10	53	60
Northwestern South Carolina:	SC-03	81	22	14	(57)	70
Greenville-Spartanburg area:	SC-04	75	23	11	54	64
Southern Charlotte suburbs and exurbs:	SC-05	75	18	13	(54)	60
Black Belt, Columbia, and Charleston:	SC-06	57	10	8	32	32
Northeastern South Carolina and the Myrtle Beach area:	SC-07	76	16	11	(52)	60
All of South Dakota:	SD-01	67	53	(10)	(47)	66
Northeastern Tennessee:	TN-01	82	38	(6)	(57)	80
Knoxville area:	TN-02	73	26	3	(51)	69
Chattanooga area:	TN-03	74	18	4	(45)	69
Southern part of Middle Tennessee:	TN-04	77	27	4	57	72
Nashville:	TN-05	51	10	2	26	41
Nashville’s eastern suburbs and northern Middle Tenn....:	TN-06	78	34	5	(57)	76
Nashville’s southern suburbs and western Middle Tenn....:	TN-07	77	24	5	55	71
West Tennessee:	TN-08	80	17	7	(60)	69
Memphis:	TN-09	47	5	2	24	21
Tyler and East Texas:	TX-01	85	38	16	67	75
Northern Houston and part of west Houston:	TX-02	65	39	5	46	55
Collin County:	TX-03	66	36	5	46	58
Northeastern Texas:	TX-04	84	45	15	68	78
Eastern edge of Dallas and rural areas to its east:	TX-05	75	36	8	52	65
Arlington and rural areas south of Dallas:	TX-06	70	31	7	48	57

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

		Non-Whites				All
		White	All	Black	Hisp.	
	West Houston: TX-07	62	30	4	37	50
	Suburbs and exurbs north of Houston: TX-08	82	50	12	64	76
	Southwest Houston: TX-09	41	13	3	19	19
	Austin to western Houston: TX-10	65	35	6	42	55
	Midland and San Angelo parts of rural west Texas: TX-11	87	63	14	67	81
	Parker and western Tarrant Counties: TX-12	73	42	8	51	66
	Texas Panhandle: TX-13	88	62	17	71	83
	Galveston area: TX-14	74	32	9	50	61
	McAllen in the Rio Grande Valley and rural counties: TX-15	65	32	5	30	42
	El Paso: TX-16	54	21	3	19	29
	College Station, Waco, and a slice of Austin suburbs: TX-17	69	35	7	49	60
	Heavily black parts of central Houston: TX-18	40	13	3	21	21
	Lubbock and rural West Texas: TX-19	84	56	13	62	76
	Downtown San Antonio: TX-20	55	26	3	26	36
	Parts of Austin and San Antonio connected by Texas H...: TX-21	63	37	4	37	56
	Southern Houston suburbs: TX-22	68	39	6	47	54
	El Paso to San Antonio: TX-23	68	37	6	36	49
	Suburbs north of Dallas and Fort Worth: TX-24	65	33	6	35	54
	Part of Austin and rural areas north toward Dallas-F...: TX-25	64	32	5	38	58
	Denton County: TX-26	71	40	7	49	64
	Corpus Christi and Victoria: TX-27	74	46	7	49	63
	Laredo area: TX-28	62	30	4	30	40
	Heavily Latino areas north and east of Downtown Houston: TX-29	52	19	3	20	27
	Downtown and South Dallas: TX-30	41	11	3	19	19
	Williamson and Bell Counties: TX-31	66	34	5	43	57
	Suburban North Dallas: TX-32	60	28	5	34	49
	Connects Downtown Fort Worth and western Dallas: TX-33	47	15	3	17	25
	Brownsville in the Rio Grande Valley and rural count...: TX-34	63	29	(4)	28	39
	San Antonio and heavily Latino areas of Austin: TX-35	47	22	3	23	33
	Southern part of East Texas: TX-36	82	47	12	60	74
	Ogden and northern Utah: UT-01	71	53	(23)	46	69
	Salt Lake City and southwestern Utah: UT-02	61	46	(16)	39	59
	Provo area and southeastern Utah: UT-03	69	50	(21)	43	67
	Southern Salt Lake County and rural areas to the south: UT-04	57	43	(15)	32	55
	Western Chesapeake Bay and exurbs of D.C. and Richmond: VA-01	68	27	8	48	57
	Virginia Beach: VA-02	62	22	8	38	52
	Hampton Roads: VA-03	56	12	8	33	34
	Richmond and Southside Virginia: VA-04	59	13	9	39	39
	Charlottesville and South-central Virginia: VA-05	66	15	7	(46)	56
	Shenandoah Valley: VA-06	69	24	8	51	63
	Suburban Richmond: VA-07	65	23	7	49	54
	Inner D.C. suburbs: VA-08	29	13	2	18	22
	Southwest Virginia: VA-09	75	26	9	(52)	72
	Loudoun County and D.C. suburbs and exurbs: VA-10	52	31	4	35	45

Table S.7 – Point Estimates of 2016 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
D.C. suburbs in Fairfax and Prince William Counties: VA-11	39	18	3	23	29
All of Vermont: VT-01	36	21	(2)	(15)	35
Northern Seattle suburbs: WA-01	43	32	(8)	32	41
Northern Puget Sound: WA-02	40	27	(7)	21	38
Vancouver and southwestern Washington: WA-03	56	41	(12)	34	54
Yakima and the Tri-Cities: WA-04	67	49	(16)	45	63
Spokane area and rural eastern Washington: WA-05	59	44	(13)	39	58
Olympic Peninsula and Tacoma: WA-06	45	30	9	29	44
Seattle: WA-07	15	8	1	7	13
Eastern Seattle suburbs and two rural counties east ...: WA-08	51	37	(10)	36	49
Southeastern Seattle and the suburbs to its south an...: WA-09	30	17	5	17	25
Olympia area and Tacoma suburbs: WA-10	46	31	9	31	44
Southeastern Wisconsin: WI-01	59	33	6	41	56
Madison area: WI-02	33	15	2	17	31
Southwestern Wisconsin: WI-03	53	34	(5)	(34)	53
Milwaukee: WI-04	33	10	3	21	23
Milwaukee suburbs: WI-05	63	41	(7)	41	61
Fox River Valley and Sheboygan: WI-06	60	37	(6)	(42)	59
Northwestern Wisconsin: WI-07	62	45	(6)	(43)	61
Northeastern Wisconsin: WI-08	61	42	(6)	(39)	60
Northern West Virginia: WV-01	74	35	(9)	(57)	72
Charleston to the Eastern Panhandle: WV-02	72	27	8	(51)	69
Southern West Virginia: WV-03	78	28	11	(58)	76
All of Wyoming: WY-01	77	63	(27)	59	76
National	59	22	7	29	49

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD

	White	Non-Whites			All
		All	Black	Hisp.	
All of Alaska: AK-01	58	48	17	46	55
Greater Mobile: AL-01	81	24	14	(76)	65
Southeastern Alabama: AL-02	84	24	15	(76)	65
Eastern Alabama: AL-03	81	23	15	(74)	66
North-central Alabama: AL-04	86	52	21	80	82
Northern Alabama: AL-05	73	31	15	62	64
Birmingham suburbs: AL-06	79	30	12	73	68
Birmingham and the Black Belt: AL-07	63	9	6	(54)	29
Northeastern Arkansas: AR-01	79	27	15	(66)	71
Greater Little Rock: AR-02	65	19	9	51	55
Northwest Arkansas: AR-03	67	49	10	53	64
Southern and western Arkansas: AR-04	79	27	15	(66)	70
Northeastern Arizona: AZ-01	55	41	19	40	49
Southeastern Arizona and eastern Tucson area: AZ-02	51	32	17	31	45
Southwestern Arizona and western Tucson area: AZ-03	44	32	13	31	36
North-central Arizona and Phoenix exurbs: AZ-04	71	58	(26)	58	69
Phoenix suburbs of Mesa and Gilbert: AZ-05	62	45	20	45	58
Scottsdale and North Phoenix: AZ-06	56	39	14	40	52
Downtown and West Phoenix: AZ-07	33	21	7	22	25
Western suburbs of Phoenix: AZ-08	63	46	19	47	58
Central Phoenix and eastern suburbs: AZ-09	44	27	12	27	38
Northeastern California: CA-01	60	46	(20)	47	58
Northern California coast: CA-02	26	19	(6)	19	25
Southern Sacramento Valley: CA-03	51	33	14	34	44
Northern Sacramento suburbs and Sierra Nevada foothills: CA-04	57	46	(19)	46	55
Wine country: CA-05	31	19	6	19	26
Sacramento: CA-06	35	21	8	23	28
Suburban Sacramento County: CA-07	48	35	13	41	43
Northern San Bernardino County and the High Desert: CA-08	64	43	21	46	56
Stockton area and eastern Contra Costa County: CA-09	49	34	12	37	41
Modesto area: CA-10	54	41	15	42	49
Contra Costa County: CA-11	29	19	6	22	24
San Francisco: CA-12	14	10	4	11	12
Oakland area: CA-13	13	7	2	9	10
San Mateo County: CA-14	25	17	7	20	21
Eastern Alameda County: CA-15	32	23	7	27	27
Fresno area: CA-16	50	34	11	37	40
Northeast San Jose area: CA-17	33	22	9	28	26
Silicon Valley: CA-18	25	17	(5)	21	22
San Jose: CA-19	36	25	9	26	29
Monterey Bay: CA-20	31	21	(7)	21	26
Southern Central Valley and part of Bakersfield: CA-21	52	41	15	43	45
Central Valley near Fresno and Tulare: CA-22	59	47	20	48	53

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

		Non-Whites				All
		White	All	Black	Hisp.	
	Bakersfield area: CA-23	66	48	24	50	59
	San Luis Obispo and Santa Barbara: CA-24	42	30	(11)	31	38
	Northern Los Angeles County suburbs: CA-25	54	37	16	39	45
	Ventura County: CA-26	42	32	(11)	32	37
	San Gabriel Valley: CA-27	38	28	12	32	32
	Burbank and Glendale: CA-28	30	23	7	27	28
	East San Fernando Valley: CA-29	33	21	8	20	24
	West San Fernando Valley: CA-30	34	24	8	24	30
	San Bernardino area: CA-31	49	35	16	39	40
	East San Gabriel Valley: CA-32	44	31	12	30	34
	Coastal West Side of Los Angeles: CA-33	33	23	8	26	30
	Downtown Los Angeles: CA-34	21	17	4	17	17
	Southwest San Bernardino County and Pomona: CA-35	44	31	11	33	34
	Eastern Riverside County: CA-36	50	34	15	35	43
	Culver City area of Los Angeles: CA-37	20	12	6	15	14
	Southeastern Los Angeles County: CA-38	46	30	10	30	33
	San Gabriel Valley and northern Orange County: CA-39	52	40	16	44	45
	East Los Angeles: CA-40	31	20	6	21	21
	Riverside and Moreno Valley: CA-41	47	33	13	35	37
	Southwestern Riverside County: CA-42	61	47	22	49	54
	Inland cities south of Los Angeles: CA-43	31	19	6	24	21
	South Los Angeles, including Compton: CA-44	32	18	6	21	20
	Irvine area of Orange County: CA-45	50	35	(16)	41	44
	Santa Ana and the Anaheim area of Orange County: CA-46	42	31	10	31	34
	Long Beach area: CA-47	43	31	10	34	36
	Coastal Orange County: CA-48	53	44	(16)	44	49
	Southern Orange and northern San Diego Counties: CA-49	47	35	(13)	37	44
	Inland San Diego County: CA-50	59	44	18	43	54
	South San Diego and Imperial County: CA-51	43	28	10	29	32
	Suburban San Diego: CA-52	39	27	10	30	35
	Eastern San Diego: CA-53	37	26	10	28	32
	Denver: CO-01	26	17	7	19	23
	Boulder and Fort Collins: CO-02	36	25	(13)	24	35
	Pueblo and the Western Slope: CO-03	56	43	(20)	41	53
	Eastern Colorado and Denver exurbs: CO-04	61	51	(23)	51	58
	Colorado Springs: CO-05	60	44	21	49	57
	Aurora and Denver's southern suburbs: CO-06	46	30	17	32	40
	Denver's northern and western suburbs: CO-07	41	32	(13)	32	38
	Hartford area: CT-01	41	25	13	31	36
	Eastern Connecticut: CT-02	46	34	22	37	45
	New Haven area: CT-03	44	26	14	35	39
	Southern Fairfield County: CT-04	39	27	15	32	35
	Northern Fairfield County and northwestern Connecticut: CT-05	48	33	16	37	45

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
All of Delaware: DE-01	48	22	12	33	41
Pensacola area: FL-01	75	40	20	62	67
Florida panhandle: FL-02	74	35	18	63	68
North Florida, including Gainesville and Ocala: FL-03	65	29	12	51	57
Part of the Jacksonville area and St. Augustine: FL-04	67	38	14	55	61
Jacksonville and Tallahassee areas: FL-05	61	15	9	47	37
Daytona area: FL-06	64	35	12	48	59
Northern Orlando suburbs: FL-07	53	31	12	36	45
Space Coast: FL-08	64	37	18	49	59
Southern suburbs of Orlando: FL-09	60	35	10	40	47
Orlando: FL-10	53	25	9	37	38
North-central Florida and The Villages: FL-11	70	41	14	56	66
Northern Tampa suburbs: FL-12	62	39	13	45	59
St. Petersburg area: FL-13	53	26	8	43	48
Tampa area: FL-14	53	29	8	38	42
Lakeland and the exurbs of Tampa and Orlando: FL-15	63	37	12	50	54
Sarasota area: FL-16	59	36	10	48	54
South-central Florida: FL-17	68	41	17	53	64
Treasure Coast and Palm Beach area: FL-18	62	32	11	48	54
Cape Coral and Fort Myers: FL-19	64	41	13	52	60
West Palm Beach and Fort Lauderdale areas: FL-20	46	14	5	33	22
Southeast Palm Beach County: FL-21	48	26	8	37	42
Boca Raton and northeast Broward County: FL-22	50	26	7	35	43
Broward County-centered: FL-23	51	32	7	40	42
Northern Miami: FL-24	44	20	7	37	24
Northwestern Miami suburbs and eastern Collier Count...: FL-25	72	57	17	60	62
Southwestern Miami area and the Florida Keys: FL-26	65	50	12	55	53
Southern Miami area and Coral Gables: FL-27	58	45	12	48	49
Coastal Georgia: GA-01	72	22	13	55	56
Southwestern Georgia: GA-02	74	14	10	55	44
Southwestern Atlanta exurbs: GA-03	76	29	18	58	63
Eastern Atlanta suburbs: GA-04	52	9	4	29	20
Downtown Atlanta: GA-05	30	4	2	17	13
Northern Atlanta suburbs: GA-06	55	25	7	27	45
Northeastern Atlanta suburbs: GA-07	62	32	8	40	47
South-central Georgia: GA-08	80	24	16	69	63
Northeastern Georgia: GA-09	82	54	25	64	77
East-Central Georgia: GA-10	75	24	14	60	60
Northwestern Atlanta suburbs and exurbs: GA-11	69	30	11	45	58
Augusta and southeastern Georgia: GA-12	77	20	14	57	57
Southwestern Atlanta suburbs: GA-13	55	11	7	34	24
Northwestern Georgia: GA-14	80	43	19	62	75
Honolulu: HI-01	43	32	14	28	35

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

		White	Non-Whites			All
			All	Black	Hisp.	
Northern Oahu and all other islands:	HI-02	41	29	(14)	25	35
	Northeastern Iowa: IA-01	53	34	(16)	(43)	52
	Southeastern Iowa: IA-02	54	35	15	46	52
	Des Moines and southwestern Iowa: IA-03	52	35	15	43	50
	Northwestern Iowa: IA-04	65	48	(22)	51	64
	Idaho Panhandle and western Boise area: ID-01	70	66	(31)	64	69
	Boise and Eastern Idaho: ID-02	63	57	(27)	54	62
	South Side Chicago and southern suburbs: IL-01	44	11	6	42	25
	South Side Chicago and southwestern suburbs: IL-02	41	9	4	34	22
	Southwestern Chicago area: IL-03	49	31	7	37	44
	Heavily Latino areas of Chicago's West and South sid...: IL-04	23	15	4	17	18
	North Side Chicago: IL-05	29	17	4	22	27
	Western Chicago suburbs: IL-06	47	28	7	36	44
	West Side and Downtown Chicago: IL-07	22	6	2	19	13
	Northwestern Chicago suburbs: IL-08	43	31	6	40	40
	Northern Chicago Suburbs: IL-09	32	15	4	23	28
	Chicago's North Shore suburbs: IL-10	39	24	6	33	35
	Southwestern Chicago suburbs: IL-11	43	26	6	37	37
	Southwestern Illinois: IL-12	63	21	13	(52)	57
	Central Illinois: IL-13	56	22	11	(44)	52
	Western Chicago exurbs: IL-14	51	35	10	42	49
	East-central and southeastern Illinois: IL-15	75	41	22	(69)	74
	North-central Illinois: IL-16	60	40	14	53	58
	Northwestern Illinois: IL-17	55	24	9	47	51
	West-central Illinois: IL-18	64	35	14	(58)	63
	Northwestern Indiana: IN-01	53	26	14	37	46
	North-central Indiana: IN-02	64	34	16	49	61
	Northeastern Indiana: IN-03	68	38	19	53	65
	West-central Indiana: IN-04	67	43	24	55	65
	Northern Indianapolis suburbs: IN-05	56	24	11	34	51
	Southeastern Indiana: IN-06	72	45	(24)	(55)	70
	Indianapolis: IN-07	43	16	8	30	36
	Southwestern Indiana: IN-08	68	36	20	(52)	66
	South-central Indiana and southern Indianapolis suburbs: IN-09	64	37	16	(46)	62
	Western Kansas: KS-01	73	61	39	65	71
	Topeka, Lawrence, and eastern rural Kansas: KS-02	59	45	27	54	58
	Suburban Kansas City area: KS-03	49	29	12	39	45
	Wichita area: KS-04	64	47	24	57	61
	Western Kentucky: KY-01	77	36	19	(64)	74
	West-central Kentucky: KY-02	72	36	14	(61)	69
	Louisville: KY-03	48	13	6	34	39
	Cincinnati suburbs and northern Kentucky border area: KY-04	68	39	11	(54)	66

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Appalachian Kentucky: KY-05	82	57	(23)	(74)	81
Greater Lexington area: KY-06	58	29	17	45	55
Suburban New Orleans: LA-01	78	42	13	65	69
New Orleans and Baton Rouge: LA-02	52	9	4	36	24
Southwestern Louisiana: LA-03	84	28	17	72	69
Northwestern Louisiana: LA-04	83	25	17	70	63
Northeastern Louisiana: LA-05	87	24	18	(78)	65
Baton Rouge suburbs: LA-06	79	29	16	67	65
Springfield area and the Berkshires: MA-01	40	26	15	27	38
Worcester area and upper Pioneer Valley: MA-02	39	26	13	25	37
Lowell area and the Merrimack Valley: MA-03	39	26	12	25	35
Western Boston suburbs and Bristol County: MA-04	35	28	11	25	35
Boston's northern and outer western suburbs: MA-05	26	18	7	16	24
North Shore: MA-06	38	27	16	25	36
Central Boston area: MA-07	16	10	5	11	14
Southern Boston suburbs: MA-08	36	22	12	25	33
Cape Cod and South Shore: MA-09	42	30	14	30	41
Eastern Shore and Baltimore exurbs: MD-01	67	25	12	(47)	60
Suburban Baltimore area: MD-02	44	13	6	27	33
Baltimore suburbs to Annapolis to D.C. suburbs: MD-03	37	15	8	20	30
Prince George's County D.C. suburbs: MD-04	41	11	4	24	20
Southern Maryland and outer Prince George's County: MD-05	48	14	6	32	30
Western Maryland and Montgomery County D.C. suburbs: MD-06	45	23	6	27	38
Baltimore and western suburbs: MD-07	34	9	6	19	20
Montgomery County: MD-08	36	17	6	19	30
Portland and Augusta: ME-01	39	28	(11)	(33)	38
Northern Maine: ME-02	54	42	(16)	(46)	54
Upper Peninsula and northern Lower Peninsula: MI-01	60	44	(14)	(49)	59
Western Michigan coast: MI-02	60	35	12	45	56
Grand Rapids area: MI-03	55	31	11	44	52
Central Michigan: MI-04	63	43	(14)	(53)	62
Flint, Saginaw, and Bay City: MI-05	55	19	9	43	48
Southwestern Michigan: MI-06	56	30	14	44	52
South-central Michigan: MI-07	60	37	13	49	58
Lansing and northwestern Detroit exurbs: MI-08	54	30	11	41	51
Northern Detroit suburbs: MI-09	48	20	10	(40)	43
Northern Macomb County and the "Thumb" region: MI-10	67	43	15	54	65
Northwestern Detroit suburbs: MI-11	52	28	9	37	48
Ann Arbor and Dearborn: MI-12	39	18	6	27	35
Western Detroit: MI-13	35	8	4	25	20
Eastern Detroit to Pontiac: MI-14	39	7	4	30	20
Southern Minnesota: MN-01	57	40	(19)	40	55

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD (*continued*)

	White	Non-Whites			All
		All	Black	Hisp.	
Southern Twin Cities suburbs: MN-02	49	32	14	36	47
Western Twin Cities suburbs: MN-03	44	25	15	28	40
St. Paul: MN-04	35	19	10	22	31
Minneapolis: MN-05	22	10	4	12	18
Northern Twin Cities exurbs: MN-06	62	45	23	(47)	60
Western Minnesota: MN-07	66	51	(26)	49	65
Iron Range: MN-08	58	46	(22)	(40)	58
St. Louis and northwestern suburbs: MO-01	27	8	6	(18)	18
Suburban St. Louis: MO-02	52	32	15	(38)	50
East-central Missouri: MO-03	70	49	29	(56)	68
West-central Missouri: MO-04	69	48	27	(58)	68
Kansas City area: MO-05	47	20	11	34	41
Northern Missouri: MO-06	66	47	31	53	65
Southwestern Missouri: MO-07	72	58	(31)	(59)	72
Southeastern Missouri: MO-08	80	58	45	(70)	79
Northeastern Mississippi: MS-01	83	22	15	(75)	66
Mississippi Delta and Jackson: MS-02	76	11	9	(66)	35
Southwestern and eastern Mississippi: MS-03	83	21	16	(75)	61
Gulf Coast: MS-04	81	28	15	73	69
All of Montana: MT-01	59	54	(30)	(53)	59
Durham and northeastern North Carolina: NC-01	67	23	17	50	46
Raleigh exurbs: NC-02	43	17	8	25	35
Eastern North Carolina coast: NC-03	73	29	15	57	62
Raleigh, Cary, Chapel Hill: NC-04	44	17	7	25	33
Winston-Salem and northwestern North Carolina: NC-05	76	36	20	58	68
Part of Greensboro and rural Piedmont: NC-06	47	16	7	30	38
Southeastern North Carolina: NC-07	70	28	14	49	59
Charlotte suburbs to Fayetteville: NC-08	66	28	14	49	53
South Charlotte to suburban Fayetteville: NC-09	65	31	13	47	54
Asheville and southwestern Piedmont: NC-10	75	37	20	59	69
Appalachian North Carolina: NC-11	58	37	14	38	56
Charlotte: NC-12	46	16	6	27	29
Greensboro and rural western Piedmont: NC-13	80	38	25	65	68
All of North Dakota: ND-01	69	51	(27)	(53)	67
Lincoln and rural eastern Nebraska: NE-01	60	45	22	47	58
Greater Omaha area: NE-02	52	29	18	36	47
Western Nebraska: NE-03	78	68	(41)	67	77
Eastern New Hampshire: NH-01	48	37	(17)	(34)	47
Western and northern New Hampshire: NH-02	46	34	(15)	(32)	46
Inner Philadelphia suburbs: NJ-01	43	24	10	32	37
Southern New Jersey coast: NJ-02	57	35	13	46	52

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Philadelphia suburbs and central Jersey Shore: NJ-03	55	31	16	41	50
Monmouth County area: NJ-04	58	40	18	48	55
Northern Bergen County and northwestern exurbs: NJ-05	52	35	20	36	48
Northern Middlesex County and northern Jersey Shore: NJ-06	49	30	12	36	42
Hunterdon County and New York City's western exurbs: NJ-07	49	34	14	37	45
Jersey City and inner New York City suburbs: NJ-08	32	23	7	25	26
Paterson and inner New York City suburbs: NJ-09	45	30	11	35	37
Newark area: NJ-10	29	12	6	19	17
Morris County area: NJ-11	50	34	14	39	47
Trenton and southern Middlesex County: NJ-12	40	20	8	27	32
Albuquerque area: NM-01	44	34	14	35	38
Southern New Mexico: NM-02	62	50	(26)	51	56
Northern New Mexico: NM-03	47	36	(19)	38	41
Las Vegas: NV-01	44	32	12	37	37
Northern Nevada: NV-02	58	49	(22)	53	55
Southern Las Vegas suburbs: NV-03	55	42	18	58	50
Northern Las Vegas suburbs and rural central Nevada: NV-04	56	39	17	52	48
Eastern Suffolk County: NY-01	56	34	19	37	52
Southern Long Island: NY-02	60	35	17	38	52
Northern Long Island: NY-03	50	30	15	32	45
Long Island: NY-04	54	30	13	37	44
Southeastern Queens: NY-05	35	13	7	21	16
Northeastern Queens: NY-06	46	27	11	29	38
Heavily Latino parts of northern Brooklyn and southw...: NY-07	23	13	5	15	18
Eastern Brooklyn: NY-08	30	9	5	17	17
Central Brooklyn: NY-09	29	10	6	20	18
West Side of Manhattan and part of southern Brooklyn: NY-10	26	14	11	15	23
Staten Island and part of southern Brooklyn: NY-11	61	38	20	45	55
East Side of Manhattan and nearby parts of Queens an...: NY-12	17	9	5	10	15
Harlem area: NY-13	19	10	4	13	12
East Bronx and northern Queens: NY-14	37	20	8	23	26
The Bronx: NY-15	29	13	4	16	14
Southern Westchester County and the northern Bronx: NY-16	37	13	7	21	24
Westchester County: NY-17	48	25	12	29	40
Lower Hudson Valley: NY-18	53	30	15	33	48
Hudson Valley: NY-19	52	27	14	33	49
Albany area: NY-20	43	21	11	28	40
North Country: NY-21	57	32	(20)	(37)	55
Binghamton and Utica-Rome: NY-22	58	30	18	(36)	56
Ithaca and southwestern New York: NY-23	57	31	(17)	(37)	56
Syracuse area: NY-24	48	20	11	(30)	46
Greater Rochester: NY-25	44	18	10	27	39
Greater Buffalo: NY-26	41	17	11	26	36

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

		White	Non-Whites		All		
			All	Black		Hisp.	
Suburbs and rural areas between Buffalo and Rochester:	NY-27	59	37	(22)	(42)	58	
	Suburban Cincinnati:	OH-01	62	18	11	(43)	52
Eastern Cincinnati and east along the Ohio River:	OH-02	60	24	16	(42)	57	
	Columbus:	OH-03	37	11	8	22	29
West-Central Ohio to North-Central Ohio:	OH-04	71	35	19	(55)	68	
	Northwestern Ohio:	OH-05	65	38	15	46	63
	Appalachian Ohio:	OH-06	74	38	(21)	(55)	73
	Canton area:	OH-07	68	31	17	(50)	66
Cincinnati suburbs and Springfield:	OH-08	70	34	18	(51)	67	
Toledo to West Side of Cleveland:	OH-09	46	16	7	29	40	
	Greater Dayton area:	OH-10	59	19	11	(41)	52
	East Cleveland to Akron:	OH-11	34	5	3	22	19
	Northern Columbus area:	OH-12	56	27	13	(38)	53
	Mahoning Valley and Akron:	OH-13	52	16	9	(35)	48
	Northeastern Ohio:	OH-14	56	29	16	(38)	55
Southern Columbus area and Athens:	OH-15	59	28	13	(38)	57	
Western suburbs of Cleveland and Akron:	OH-16	59	34	(13)	(42)	57	
	Greater Tulsa area:	OK-01	66	45	13	62	62
	Eastern Oklahoma:	OK-02	80	62	21	70	78
	Northwestern Oklahoma:	OK-03	79	59	21	73	76
	Southern Oklahoma:	OK-04	70	49	14	66	67
	Oklahoma City area:	OK-05	58	33	10	51	53
	Western Portland suburbs and North Coast:	OR-01	37	28	(9)	29	35
Oregon east of the Cascades and part of southern Oregon:	OR-02	58	49	(18)	48	57	
	Eastern Portland area:	OR-03	26	17	6	19	24
Southern Willamette Valley and southern coast:	OR-04	49	43	(14)	40	48	
Salem area and southern Portland suburbs:	OR-05	48	34	(14)	32	45	
	Philadelphia's Bucks County suburbs:	PA-01	51	25	12	33	47
	Northeast Philadelphia:	PA-02	40	17	9	23	30
West, Northwest, and Center City Philadelphia:	PA-03	19	4	2	11	9	
Philadelphia's Montgomery County suburbs:	PA-04	43	18	10	26	38	
Philadelphia's Delaware County suburbs and South Phi...:	PA-05	44	14	8	(26)	34	
Philadelphia's Chester County suburbs and Reading:	PA-06	48	24	9	27	43	
Lehigh Valley and southern Monroe County:	PA-07	52	29	11	34	48	
Scranton/Wilkes-Barre and northeastern Pennsylvania:	PA-08	56	27	12	33	52	
Coal Region between Wilkes-Barre and Harrisburg:	PA-09	67	39	19	47	66	
	Harrisburg and York:	PA-10	57	26	13	39	52
Lancaster and southern York Counties:	PA-11	64	38	18	46	61	
Williamsport area and Northern Tier:	PA-12	70	39	(21)	(50)	69	
	South-central Pennsylvania:	PA-13	74	43	27	(57)	73
Southern Pittsburgh exurbs and southwestern Pennsylv...:	PA-14	65	30	19	(46)	64	
	North-central Pennsylvania:	PA-15	73	39	(24)	(54)	72
	Erie and northwestern Pennsylvania:	PA-16	61	28	17	(41)	60

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD (*continued*)

	White	Non-Whites			All
		All	Black	Hisp.	
Pittsburgh suburbs and Beaver County: PA-17	51	22	12	(35)	49
Pittsburgh: PA-18	40	13	9	(25)	35
Part of Providence and eastern Rhode Island: RI-01	38	25	10	31	35
Part of Providence and western Rhode Island: RI-02	45	32	13	35	43
Charleston and coastal South Carolina: SC-01	63	23	8	53	53
Augusta and Columbia suburbs: SC-02	70	20	9	59	56
Northwestern South Carolina: SC-03	78	28	17	(70)	69
Greenville-Spartanburg area: SC-04	71	30	12	60	61
Southern Charlotte suburbs and exurbs: SC-05	73	22	12	64	59
Black Belt, Columbia, and Charleston: SC-06	59	9	6	(51)	32
Northeastern South Carolina and the Myrtle Beach area: SC-07	73	18	10	(63)	60
All of South Dakota: SD-01	64	55	(22)	(57)	64
Northeastern Tennessee: TN-01	79	49	(15)	(69)	78
Knoxville area: TN-02	68	34	10	(59)	65
Chattanooga area: TN-03	71	30	10	(61)	67
Southern part of Middle Tennessee: TN-04	74	35	11	64	69
Nashville: TN-05	48	13	3	36	38
Nashville's eastern suburbs and northern Middle Tenn...: TN-06	76	46	16	(66)	74
Nashville's southern suburbs and western Middle Tenn...: TN-07	74	35	11	62	68
West Tennessee: TN-08	77	26	13	(66)	67
Memphis: TN-09	45	7	3	34	20
Tyler and East Texas: TX-01	82	40	20	68	73
Northern Houston and part of west Houston: TX-02	59	38	8	49	51
Collin County: TX-03	59	34	11	41	51
Northeastern Texas: TX-04	82	46	20	67	75
Eastern edge of Dallas and rural areas to its east: TX-05	71	37	13	54	62
Arlington and rural areas south of Dallas: TX-06	64	32	8	52	52
West Houston: TX-07	56	32	6	40	46
Suburbs and exurbs north of Houston: TX-08	77	52	20	65	72
Southwest Houston: TX-09	45	18	5	28	24
Austin to western Houston: TX-10	61	34	8	41	51
Midland and San Angelo parts of rural west Texas: TX-11	85	67	25	72	80
Parker and western Tarrant Counties: TX-12	68	43	12	54	62
Texas Panhandle: TX-13	85	63	25	71	81
Galveston area: TX-14	72	36	14	55	60
McAllen in the Rio Grande Valley and rural counties ...: TX-15	65	43	11	43	49
El Paso: TX-16	47	28	6	28	33
College Station, Waco, and a slice of Austin suburbs: TX-17	64	34	10	47	56
Heavily black parts of central Houston: TX-18	40	17	4	30	23
Lubbock and rural West Texas: TX-19	80	57	21	64	73
Downtown San Antonio: TX-20	49	28	8	29	35
Parts of Austin and San Antonio connected by Texas H...: TX-21	57	38	8	40	52

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD *(continued)*

		White	Non-Whites		All
			All	Black Hisp.	
	Southern Houston suburbs: TX-22	64	38	9 48	51
	El Paso to San Antonio: TX-23	66	43	10 44	51
	Suburbs north of Dallas and Fort Worth: TX-24	56	33	9 41	47
	Part of Austin and rural areas north toward Dallas-F...: TX-25	60	38	9 48	55
	Denton County: TX-26	64	41	9 53	57
	Corpus Christi and Victoria: TX-27	72	47	14 50	62
	Laredo area: TX-28	61	43	9 45	48
	Heavily Latino areas north and east of Downtown Houston: TX-29	50	28	6 33	33
	Downtown and South Dallas: TX-30	38	12	4 24	19
	Williamson and Bell Counties: TX-31	59	36	9 46	52
	Suburban North Dallas: TX-32	53	30	8 38	45
	Connects Downtown Fort Worth and western Dallas: TX-33	40	20	6 24	26
	Brownsville in the Rio Grande Valley and rural count...: TX-34	60	44	(8) 44	48
	San Antonio and heavily Latino areas of Austin: TX-35	41	24	6 26	31
	Southern part of East Texas: TX-36	79	50	18 63	73
	Ogden and northern Utah: UT-01	69	58	(33) 56	67
	Salt Lake City and southwestern Utah: UT-02	60	50	(25) 46	58
	Provo area and southeastern Utah: UT-03	65	52	(28) 50	63
	Southern Salt Lake County and rural areas to the south: UT-04	57	48	(25) 44	55
	Western Chesapeake Bay and exurbs of D.C. and Richmond: VA-01	62	30	10 46	52
	Virginia Beach: VA-02	55	26	10 42	48
	Hampton Roads: VA-03	51	12	6 36	32
	Richmond and Southside Virginia: VA-04	53	16	10 39	38
	Charlottesville and South-central Virginia: VA-05	63	22	13 (45)	54
	Shenandoah Valley: VA-06	66	30	13 52	61
	Suburban Richmond: VA-07	60	25	10 42	50
	Inner D.C. suburbs: VA-08	27	14	3 19	22
	Southwest Virginia: VA-09	74	33	16 (56)	71
	Loudoun County and D.C. suburbs and exurbs: VA-10	46	30	7 31	40
	D.C. suburbs in Fairfax and Prince William Counties: VA-11	37	20	4 24	29
	All of Vermont: VT-01	32	22	(8) (24)	32
	Northern Seattle suburbs: WA-01	41	32	(12) 35	39
	Northern Puget Sound: WA-02	38	30	13 28	36
	Vancouver and southwestern Washington: WA-03	53	43	(20) 41	52
	Yakima and the Tri-Cities: WA-04	63	51	(21) 50	59
	Spokane area and rural eastern Washington: WA-05	56	43	(18) 42	55
	Olympic Peninsula and Tacoma: WA-06	43	31	11 31	41
	Seattle: WA-07	14	9	2 11	13
	Eastern Seattle suburbs and two rural counties east ...: WA-08	49	38	(15) 38	47
	Southeastern Seattle and the suburbs to its south an...: WA-09	29	20	6 21	25
	Olympia area and Tacoma suburbs: WA-10	45	31	12 33	42
	Southeastern Wisconsin: WI-01	58	37	13 43	55

Table S.8 – Point Estimates of 2020 Two-Party Vote for the Republican Presidential Candidate, by Race by CD (*continued*)

		White	Non-Whites			All
			All	Black	Hisp.	
	Madison area: WI-02	32	16	6	20	30
	Southwestern Wisconsin: WI-03	53	38	(13)	(42)	53
	Milwaukee: WI-04	32	12	5	21	23
	Milwaukee suburbs: WI-05	60	42	(16)	49	58
	Fox River Valley and Sheboygan: WI-06	59	39	(15)	43	58
	Northwestern Wisconsin: WI-07	61	46	(20)	(46)	60
	Northeastern Wisconsin: WI-08	59	44	(16)	46	58
	Northern West Virginia: WV-01	70	41	(15)	(54)	69
	Charleston to the Eastern Panhandle: WV-02	69	35	14	(46)	67
	Southern West Virginia: WV-03	76	40	20	(58)	74
	All of Wyoming: WY-01	73	66	(34)	67	73
	National	56	28	10	38	48

4 TURNOUT COMPOSITION ESTIMATES

Here we provide the estimates of the size of the group in the electorate. In our estimation, we ensure that the racial composition of the Voting Age Population (VAP) in the CD matches the ACS. We also ensure that the total VAP turnout rate at the district level is consistent with the true turnout. Tables S.9 and S.9 returns the population size at the district level, transformed as a percent of the entire electorate in the geography. Table S.11 and S.12 aggregates the counts to the state level.

Table S.9 – Composition of 2016 Electorate, by Race by CD

		White	Non-Whites			All
			All	Black	Hisp.	
	All of Alaska: AK-01	76	24	2	3	100
	Greater Mobile: AL-01	73	27	22	2	100
	Southeastern Alabama: AL-02	71	29	26	1	100
	Eastern Alabama: AL-03	74	26	22	1	100
	North-central Alabama: AL-04	89	11	5	3	100
	Northern Alabama: AL-05	81	19	14	2	100
	Birmingham suburbs: AL-06	81	19	13	3	100
	Birmingham and the Black Belt: AL-07	37	63	59	2	100
	Northeastern Arkansas: AR-01	85	15	12	1	100

Table S.9 – Composition of 2016 Electorate, by Race by CD *(continued)*

		White	Non-Whites			All
			All	Black	Hisp.	
	Greater Little Rock: AR-02	78	22	17	2	100
	Northwest Arkansas: AR-03	86	14	2	7	100
	Southern and western Arkansas: AR-04	81	19	15	2	100
	Northeastern Arizona: AZ-01	69	31	2	15	100
	Southeastern Arizona and eastern Tucson area: AZ-02	76	24	2	16	100
	Southwestern Arizona and western Tucson area: AZ-03	48	52	4	36	100
	North-central Arizona and Phoenix exurbs: AZ-04	87	13	1	9	100
	Phoenix suburbs of Mesa and Gilbert: AZ-05	83	17	2	10	100
	Scottsdale and North Phoenix: AZ-06	85	15	2	9	100
	Downtown and West Phoenix: AZ-07	46	54	9	31	100
	Western suburbs of Phoenix: AZ-08	83	17	3	11	100
	Central Phoenix and eastern suburbs: AZ-09	74	26	3	15	100
	Northeastern California: CA-01	88	12	1	6	100
	Northern California coast: CA-02	83	17	1	8	100
	Southern Sacramento Valley: CA-03	65	35	6	16	100
	Northern Sacramento suburbs and Sierra Nevada foothi...: CA-04	86	14	1	7	100
	Wine country: CA-05	65	35	5	15	100
	Sacramento: CA-06	55	45	11	17	100
	Suburban Sacramento County: CA-07	70	30	6	10	100
	Northern San Bernardino County and the High Desert: CA-08	65	35	6	21	100
	Stockton area and eastern Contra Costa County: CA-09	52	48	8	23	100
	Modesto area: CA-10	62	38	3	24	100
	Contra Costa County: CA-11	59	41	7	15	100
	San Francisco: CA-12	52	48	5	9	100
	Oakland area: CA-13	42	58	17	13	100
	San Mateo County: CA-14	48	52	3	15	100
	Eastern Alameda County: CA-15	49	51	6	15	100
	Fresno area: CA-16	41	59	6	34	100
	Northeast San Jose area: CA-17	42	58	3	13	100
	Silicon Valley: CA-18	67	33	2	9	100
	San Jose: CA-19	38	62	3	26	100
	Monterey Bay: CA-20	53	47	2	30	100
	Southern Central Valley and part of Bakersfield: CA-21	37	63	5	41	100
	Central Valley near Fresno and Tulare: CA-22	58	42	3	27	100
	Bakersfield area: CA-23	65	35	5	21	100
	San Luis Obispo and Santa Barbara: CA-24	70	30	2	20	100
	Northern Los Angeles County suburbs: CA-25	54	46	6	24	100
	Ventura County: CA-26	60	40	2	28	100
	San Gabriel Valley: CA-27	39	61	5	20	100
	Burbank and Glendale: CA-28	68	32	2	15	100
	East San Fernando Valley: CA-29	32	68	4	42	100
	West San Fernando Valley: CA-30	65	35	4	16	100
	San Bernardino area: CA-31	41	59	10	34	100
	East San Gabriel Valley: CA-32	24	76	3	42	100
	Coastal West Side of Los Angeles: CA-33	73	27	3	9	100

Table S.9 – Composition of 2016 Electorate, by Race by CD *(continued)*

		White	Non-Whites			All
			All	Black	Hisp.	
	Downtown Los Angeles: CA-34	17	83	5	49	100
	Southwest San Bernardino County and Pomona: CA-35	24	76	7	43	100
	Eastern Riverside County: CA-36	63	37	4	23	100
	Culver City area of Los Angeles: CA-37	33	67	23	24	100
	Southeastern Los Angeles County: CA-38	24	76	3	42	100
	San Gabriel Valley and northern Orange County: CA-39	44	56	2	24	100
	East Los Angeles: CA-40	9	91	6	61	100
	Riverside and Moreno Valley: CA-41	36	64	9	35	100
	Southwestern Riverside County: CA-42	57	43	5	23	100
	Inland cities south of Los Angeles: CA-43	22	78	22	31	100
	South Los Angeles, including Compton: CA-44	12	88	16	49	100
	Irvine area of Orange County: CA-45	66	34	2	12	100
	Santa Ana and the Anaheim area of Orange County: CA-46	34	66	2	40	100
	Long Beach area: CA-47	48	52	7	22	100
	Coastal Orange County: CA-48	69	31	1	12	100
	Southern Orange and northern San Diego Counties: CA-49	75	25	2	14	100
	Inland San Diego County: CA-50	71	29	2	18	100
	South San Diego and Imperial County: CA-51	23	77	7	50	100
	Suburban San Diego: CA-52	73	27	3	9	100
	Eastern San Diego: CA-53	55	45	7	22	100
	Denver: CO-01	67	33	6	20	100
	Boulder and Fort Collins: CO-02	89	11	0	6	100
	Pueblo and the Western Slope: CO-03	79	21	0	16	100
	Eastern Colorado and Denver exurbs: CO-04	80	20	1	15	100
	Colorado Springs: CO-05	80	20	4	10	100
	Aurora and Denver's southern suburbs: CO-06	71	29	7	13	100
	Denver's northern and western suburbs: CO-07	73	27	1	21	100
	Hartford area: CT-01	71	29	13	9	100
	Eastern Connecticut: CT-02	91	9	2	3	100
	New Haven area: CT-03	76	24	11	7	100
	Southern Fairfield County: CT-04	70	30	11	13	100
	Northern Fairfield County and northwestern Connecticut: CT-05	80	20	4	9	100
	All of Delaware: DE-01	75	25	16	5	100
	Pensacola area: FL-01	79	21	11	3	100
	Florida panhandle: FL-02	83	17	11	3	100
	North Florida, including Gainesville and Ocala: FL-03	77	23	13	5	100
	Part of the Jacksonville area and St. Augustine: FL-04	82	18	7	4	100
	Jacksonville and Tallahassee areas: FL-05	48	52	44	4	100
	Daytona area: FL-06	84	16	7	6	100
	Northern Orlando suburbs: FL-07	67	33	8	16	100
	Space Coast: FL-08	84	16	7	5	100
	Southern suburbs of Orlando: FL-09	52	48	11	28	100
	Orlando: FL-10	47	53	23	18	100
	North-central Florida and The Villages: FL-11	89	11	4	5	100

Table S.9 – Composition of 2016 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Northern Tampa suburbs: FL-12	87	13	3	6	100
St. Petersburg area: FL-13	83	17	8	5	100
Tampa area: FL-14	55	45	15	21	100
Lakeland and the exurbs of Tampa and Orlando: FL-15	70	30	11	12	100
Sarasota area: FL-16	84	16	5	7	100
South-central Florida: FL-17	87	13	4	6	100
Treasure Coast and Palm Beach area: FL-18	79	21	9	8	100
Cape Coral and Fort Myers: FL-19	84	16	4	9	100
West Palm Beach and Fort Lauderdale areas: FL-20	28	72	48	18	100
Southeast Palm Beach County: FL-21	73	27	10	13	100
Boca Raton and northeast Broward County: FL-22	72	28	10	13	100
Broward County-centered: FL-23	51	49	11	29	100
Northern Miami: FL-24	17	83	44	33	100
Northwestern Miami suburbs and eastern Collier Count...: FL-25	30	70	4	61	100
Southwestern Miami area and the Florida Keys: FL-26	23	77	9	63	100
Southern Miami area and Coral Gables: FL-27	27	73	4	63	100
Coastal Georgia: GA-01	70	30	24	3	100
Southwestern Georgia: GA-02	51	49	44	2	100
Southwestern Atlanta exurbs: GA-03	75	25	20	2	100
Eastern Atlanta suburbs: GA-04	31	69	56	4	100
Downtown Atlanta: GA-05	35	65	56	3	100
Northern Atlanta suburbs: GA-06	67	33	11	8	100
Northeastern Atlanta suburbs: GA-07	56	44	17	11	100
South-central Georgia: GA-08	70	30	25	2	100
Northeastern Georgia: GA-09	88	12	5	5	100
East-Central Georgia: GA-10	74	26	21	2	100
Northwestern Atlanta suburbs and exurbs: GA-11	74	26	12	7	100
Augusta and southeastern Georgia: GA-12	66	34	28	2	100
Southwestern Atlanta suburbs: GA-13	33	67	54	7	100
Northwestern Georgia: GA-14	86	14	7	4	100
Honolulu: HI-01	33	67	2	7	100
Northern Oahu and all other islands: HI-02	54	46	2	7	100
Northeastern Iowa: IA-01	94	6	2	2	100
Southeastern Iowa: IA-02	91	9	3	3	100
Des Moines and southwestern Iowa: IA-03	91	9	2	3	100
Northwestern Iowa: IA-04	93	7	1	3	100
Idaho Panhandle and western Boise area: ID-01	93	7	0	3	100
Boise and Eastern Idaho: ID-02	91	9	0	5	100
South Side Chicago and southern suburbs: IL-01	40	60	51	5	100
South Side Chicago and southwestern suburbs: IL-02	37	63	52	8	100
Southwestern Chicago area: IL-03	72	28	4	16	100
Heavily Latino areas of Chicago's West and South sid...: IL-04	34	66	4	43	100
North Side Chicago: IL-05	80	20	2	11	100
Western Chicago suburbs: IL-06	84	16	2	6	100

Table S.9 – Composition of 2016 Electorate, by Race by CD (*continued*)

	White	Non-Whites			All
		All	Black	Hisp.	
West Side and Downtown Chicago: IL-07	38	62	44	9	100
Northwestern Chicago suburbs: IL-08	73	27	4	13	100
Northern Chicago Suburbs: IL-09	76	24	7	7	100
Chicago's North Shore suburbs: IL-10	75	25	5	11	100
Southwestern Chicago suburbs: IL-11	66	34	9	14	100
Southwestern Illinois: IL-12	85	15	12	1	100
Central Illinois: IL-13	89	11	7	1	100
Western Chicago exurbs: IL-14	87	13	2	7	100
East-central and southeastern Illinois: IL-15	95	5	3	1	100
North-central Illinois: IL-16	92	8	2	4	100
Northwestern Illinois: IL-17	87	13	7	4	100
West-central Illinois: IL-18	94	6	2	1	100
Northwestern Indiana: IN-01	75	25	13	8	100
North-central Indiana: IN-02	91	9	5	3	100
Northeastern Indiana: IN-03	92	8	4	2	100
West-central Indiana: IN-04	93	7	3	2	100
Northern Indianapolis suburbs: IN-05	88	12	7	2	100
Southeastern Indiana: IN-06	96	4	2	1	100
Indianapolis: IN-07	74	26	19	4	100
Southwestern Indiana: IN-08	95	5	3	1	100
South-central Indiana and southern Indianapolis suburbs: IN-09	95	5	2	1	100
Western Kansas: KS-01	91	9	2	4	100
Topeka, Lawrence, and eastern rural Kansas: KS-02	92	8	2	2	100
Suburban Kansas City area: KS-03	83	17	6	6	100
Wichita area: KS-04	89	11	3	4	100
Western Kentucky: KY-01	92	8	5	1	100
West-central Kentucky: KY-02	92	8	4	2	100
Louisville: KY-03	75	25	18	4	100
Cincinnati suburbs and northern Kentucky border area: KY-04	94	6	2	2	100
Appalachian Kentucky: KY-05	98	2	1	0	100
Greater Lexington area: KY-06	88	12	6	3	100
Suburban New Orleans: LA-01	79	21	10	7	100
New Orleans and Baton Rouge: LA-02	34	66	58	4	100
Southwestern Louisiana: LA-03	74	26	21	3	100
Northwestern Louisiana: LA-04	65	35	30	2	100
Northeastern Louisiana: LA-05	66	34	31	2	100
Baton Rouge suburbs: LA-06	74	26	19	3	100
Springfield area and the Berkshires: MA-01	83	17	4	10	100
Worcester area and upper Pioneer Valley: MA-02	88	12	3	4	100
Lowell area and the Merrimack Valley: MA-03	76	24	2	12	100
Western Boston suburbs and Bristol County: MA-04	90	10	2	2	100
Boston's northern and outer western suburbs: MA-05	80	20	3	5	100
North Shore: MA-06	87	13	2	5	100
Central Boston area: MA-07	56	44	19	12	100

Table S.9 – Composition of 2016 Electorate, by Race by CD (*continued*)

	White	Non-Whites			All
		All	Black	Hisp.	
Southern Boston suburbs: MA-08	81	19	8	3	100
Cape Cod and South Shore: MA-09	92	8	2	2	100
Eastern Shore and Baltimore exurbs: MD-01	88	12	8	2	100
Suburban Baltimore area: MD-02	65	35	27	4	100
Baltimore suburbs to Annapolis to D.C. suburbs: MD-03	71	29	17	5	100
Prince George's County D.C. suburbs: MD-04	30	70	51	9	100
Southern Maryland and outer Prince George's County: MD-05	51	49	37	6	100
Western Maryland and Montgomery County D.C. suburbs: MD-06	71	29	9	9	100
Baltimore and western suburbs: MD-07	41	59	51	2	100
Montgomery County: MD-08	66	34	10	10	100
Portland and Augusta: ME-01	98	2	1	1	100
Northern Maine: ME-02	97	3	0	0	100
Upper Peninsula and northern Lower Peninsula: MI-01	96	4	1	1	100
Western Michigan coast: MI-02	88	12	4	5	100
Grand Rapids area: MI-03	89	11	5	3	100
Central Michigan: MI-04	96	4	1	1	100
Flint, Saginaw, and Bay City: MI-05	83	17	13	2	100
Southwestern Michigan: MI-06	90	10	5	2	100
South-central Michigan: MI-07	94	6	3	2	100
Lansing and northwestern Detroit exurbs: MI-08	89	11	4	3	100
Northern Detroit suburbs: MI-09	88	12	8	1	100
Northern Macomb County and the "Thumb" region: MI-10	95	5	2	1	100
Northwestern Detroit suburbs: MI-11	86	14	4	2	100
Ann Arbor and Dearborn: MI-12	85	15	7	3	100
Western Detroit: MI-13	47	53	46	4	100
Eastern Detroit to Pontiac: MI-14	39	61	54	3	100
Southern Minnesota: MN-01	92	8	1	4	100
Southern Twin Cities suburbs: MN-02	88	12	3	3	100
Western Twin Cities suburbs: MN-03	86	14	5	2	100
St. Paul: MN-04	79	21	7	4	100
Minneapolis: MN-05	74	26	12	5	100
Northern Twin Cities exurbs: MN-06	95	5	1	1	100
Western Minnesota: MN-07	95	5	1	2	100
Iron Range: MN-08	96	4	1	1	100
St. Louis and northwestern suburbs: MO-01	54	46	41	2	100
Suburban St. Louis: MO-02	91	9	3	2	100
East-central Missouri: MO-03	95	5	2	1	100
West-central Missouri: MO-04	93	7	3	2	100
Kansas City area: MO-05	77	23	15	4	100
Northern Missouri: MO-06	93	7	3	2	100
Southwestern Missouri: MO-07	95	5	1	2	100
Southeastern Missouri: MO-08	95	5	3	1	100
Northeastern Mississippi: MS-01	74	26	23	1	100

Table S.9 – Composition of 2016 Electorate, by Race by CD (*continued*)

	White	Non-Whites			All
		All	Black	Hisp.	
Mississippi Delta and Jackson: MS-02	37	63	61	1	100
Southwestern and eastern Mississippi: MS-03	67	33	30	1	100
Gulf Coast: MS-04	78	22	18	2	100
All of Montana: MT-01	95	5	0	1	100
Durham and northeastern North Carolina: NC-01	49	51	43	5	100
Raleigh exurbs: NC-02	72	28	18	6	100
Eastern North Carolina coast: NC-03	76	24	19	3	100
Raleigh, Cary, Chapel Hill: NC-04	61	39	21	7	100
Winston-Salem and northwestern North Carolina: NC-05	81	19	13	5	100
Part of Greensboro and rural Piedmont: NC-06	73	27	19	5	100
Southeastern North Carolina: NC-07	77	23	17	4	100
Charlotte suburbs to Fayetteville: NC-08	71	29	21	5	100
South Charlotte to suburban Fayetteville: NC-09	69	31	18	4	100
Asheville and southwestern Piedmont: NC-10	84	16	10	3	100
Appalachian North Carolina: NC-11	93	7	3	2	100
Charlotte: NC-12	48	52	37	9	100
Greensboro and rural western Piedmont: NC-13	72	28	20	5	100
All of North Dakota: ND-01	95	5	1	1	100
Lincoln and rural eastern Nebraska: NE-01	92	8	1	3	100
Greater Omaha area: NE-02	85	15	6	5	100
Western Nebraska: NE-03	94	6	1	4	100
Eastern New Hampshire: NH-01	96	4	1	1	100
Western and northern New Hampshire: NH-02	95	5	1	1	100
Inner Philadelphia suburbs: NJ-01	70	30	13	8	100
Southern New Jersey coast: NJ-02	75	25	9	10	100
Philadelphia suburbs and central Jersey Shore: NJ-03	83	17	8	4	100
Monmouth County area: NJ-04	84	16	6	6	100
Northern Bergen County and northwestern exurbs: NJ-05	75	25	4	11	100
Northern Middlesex County and northern Jersey Shore: NJ-06	62	38	8	14	100
Hunterdon County and New York City's western exurbs: NJ-07	78	22	4	8	100
Jersey City and inner New York City suburbs: NJ-08	36	64	9	39	100
Paterson and inner New York City suburbs: NJ-09	49	51	8	25	100
Newark area: NJ-10	26	74	48	15	100
Morris County area: NJ-11	79	21	3	8	100
Trenton and southern Middlesex County: NJ-12	59	41	14	12	100
Albuquerque area: NM-01	53	47	2	34	100
Southern New Mexico: NM-02	58	42	2	33	100
Northern New Mexico: NM-03	51	49	1	31	100
Las Vegas: NV-01	54	46	8	25	100
Northern Nevada: NV-02	77	23	1	12	100
Southern Las Vegas suburbs: NV-03	72	28	5	11	100
Northern Las Vegas suburbs and rural central Nevada: NV-04	63	37	10	18	100

Table S.9 – Composition of 2016 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Eastern Suffolk County: NY-01	85	15	4	8	100
Southern Long Island: NY-02	71	29	8	14	100
Northern Long Island: NY-03	76	24	3	6	100
Long Island: NY-04	60	40	14	13	100
Southeastern Queens: NY-05	16	84	53	14	100
Northeastern Queens: NY-06	57	43	6	15	100
Heavily Latino parts of northern Brooklyn and southw...: NY-07	40	60	12	27	100
Eastern Brooklyn: NY-08	33	67	48	11	100
Central Brooklyn: NY-09	39	61	46	8	100
West Side of Manhattan and part of southern Brooklyn: NY-10	75	25	4	9	100
Staten Island and part of southern Brooklyn: NY-11	76	24	7	9	100
East Side of Manhattan and nearby parts of Queens an...: NY-12	76	24	5	9	100
Harlem area: NY-13	16	84	27	34	100
East Bronx and northern Queens: NY-14	36	64	13	32	100
The Bronx: NY-15	3	97	31	41	100
Southern Westchester County and the northern Bronx: NY-16	45	55	31	15	100
Westchester County: NY-17	63	37	10	14	100
Lower Hudson Valley: NY-18	76	24	8	9	100
Hudson Valley: NY-19	91	9	3	3	100
Albany area: NY-20	86	14	7	3	100
North Country: NY-21	95	5	2	1	100
Binghamton and Utica-Rome: NY-22	94	6	3	2	100
Ithaca and southwestern New York: NY-23	94	6	2	2	100
Syracuse area: NY-24	90	10	6	2	100
Greater Rochester: NY-25	78	22	13	4	100
Greater Buffalo: NY-26	78	22	15	3	100
Suburbs and rural areas between Buffalo and Rochester: NY-27	96	4	2	1	100
Suburban Cincinnati: OH-01	78	22	18	1	100
Eastern Cincinnati and east along the Ohio River: OH-02	91	9	6	1	100
Columbus: OH-03	67	33	25	3	100
West-Central Ohio to North-Central Ohio: OH-04	93	7	4	2	100
Northwestern Ohio: OH-05	94	6	2	2	100
Appalachian Ohio: OH-06	97	3	2	1	100
Canton area: OH-07	95	5	3	1	100
Cincinnati suburbs and Springfield: OH-08	93	7	4	1	100
Toledo to West Side of Cleveland: OH-09	81	19	11	5	100
Greater Dayton area: OH-10	84	16	13	1	100
East Cleveland to Akron: OH-11	47	53	47	3	100
Northern Columbus area: OH-12	92	8	3	1	100
Mahoning Valley and Akron: OH-13	89	11	8	2	100
Northeastern Ohio: OH-14	94	6	3	1	100
Southern Columbus area and Athens: OH-15	94	6	3	1	100
Western suburbs of Cleveland and Akron: OH-16	96	4	2	1	100
Greater Tulsa area: OK-01	83	17	5	4	100
Eastern Oklahoma: OK-02	86	14	2	2	100

Table S.9 – Composition of 2016 Electorate, by Race by CD *(continued)*

		White	Non-Whites			All
			All	Black	Hisp.	
	Northwestern Oklahoma: OK-03	89	11	2	3	100
	Southern Oklahoma: OK-04	87	13	3	3	100
	Oklahoma City area: OK-05	81	19	6	5	100
	Western Portland suburbs and North Coast: OR-01	84	16	1	7	100
	Oregon east of the Cascades and part of southern Oregon: OR-02	91	9	0	6	100
	Eastern Portland area: OR-03	83	17	4	5	100
	Southern Willamette Valley and southern coast: OR-04	93	7	0	3	100
	Salem area and southern Portland suburbs: OR-05	88	12	1	7	100
	Philadelphia along the Delaware River and City of Ch...: PA-01	49	51	32	9	100
	Philadelphia City center and Chestnut Hill sububrs: PA-02	33	67	58	3	100
	Erie and northwestern Pennsylvania: PA-03	95	5	3	1	100
	Harrisburg and York: PA-04	89	11	5	3	100
	North-central Pennsylvania through Centre County: PA-05	95	5	2	1	100
	Philadelphia's Chester County suburbs and Reading su...: PA-06	89	11	3	3	100
	Philadelphia's Delaware and Montgomery County suburbs: PA-07	89	11	5	2	100
	Philadelphia's Bucks County suburbs: PA-08	90	10	3	3	100
	South-central Pennsylvania: PA-09	96	4	2	1	100
	Williamsport area and northeastern Pennsylvania: PA-10	95	5	3	1	100
	Coal Region between Wilkes-Barre and Harrisburg: PA-11	92	8	4	2	100
	Beaver County through Pittsburgh's North Hills to Jo...: PA-12	95	5	2	1	100
	Northeast Philadelphia and Norristown: PA-13	68	32	15	6	100
	Pittsburgh: PA-14	78	22	17	1	100
	Lehigh Valley stretched to Susquehanna River: PA-15	84	16	3	8	100
	Reading, Lancaster, and Coatesville: PA-16	82	18	5	8	100
	Scranton/Wilkes-Barre and the Slate Belt: PA-17	90	10	4	3	100
	Pittsburgh's South and West Hills, Washington, Westm...: PA-18	96	4	2	1	100
	Part of Providence and eastern Rhode Island: RI-01	83	17	4	7	100
	Part of Providence and western Rhode Island: RI-02	88	12	2	5	100
	Charleston and coastal South Carolina: SC-01	79	21	16	2	100
	Augusta and Columbia suburbs: SC-02	75	25	20	2	100
	Northwestern South Carolina: SC-03	81	19	16	2	100
	Greenville-Spartanburg area: SC-04	78	22	16	3	100
	Southern Charlotte suburbs and exurbs: SC-05	73	27	24	2	100
	Black Belt, Columbia, and Charleston: SC-06	45	55	51	2	100
	Northeastern South Carolina and the Myrtle Beach area: SC-07	74	26	23	2	100
	All of South Dakota: SD-01	93	7	1	1	100
	Northeastern Tennessee: TN-01	96	4	2	1	100
	Knoxville area: TN-02	92	8	4	2	100
	Chattanooga area: TN-03	90	10	7	2	100
	Southern part of Middle Tennessee: TN-04	90	10	6	2	100
	Nashville: TN-05	74	26	19	4	100
	Nashville's eastern suburbs and northern Middle Tenn...: TN-06	94	6	3	2	100
	Nashville's southern suburbs and western Middle Tenn...: TN-07	89	11	7	2	100

Table S.9 – Composition of 2016 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
West Tennessee: TN-08	82	18	15	1	100
Memphis: TN-09	37	63	57	4	100
Tyler and East Texas: TX-01	78	22	13	7	100
Northern Houston and part of west Houston: TX-02	62	38	9	19	100
Collin County: TX-03	72	28	8	9	100
Northeastern Texas: TX-04	84	16	7	6	100
Eastern edge of Dallas and rural areas to its east: TX-05	74	26	10	13	100
Arlington and rural areas south of Dallas: TX-06	65	35	16	12	100
West Houston: TX-07	61	39	11	19	100
Suburbs and exurbs north of Houston: TX-08	80	20	6	11	100
Southwest Houston: TX-09	21	79	38	28	100
Austin to western Houston: TX-10	67	33	8	16	100
Midland and San Angelo parts of rural west Texas: TX-11	74	26	3	19	100
Parker and western Tarrant Counties: TX-12	78	22	6	11	100
Texas Panhandle: TX-13	80	20	4	13	100
Galveston area: TX-14	67	33	16	13	100
McAllen in the Rio Grande Valley and rural counties: TX-15	29	71	2	60	100
El Paso: TX-16	24	76	4	65	100
College Station, Waco, and a slice of Austin suburbs: TX-17	73	27	10	13	100
Heavily black parts of central Houston: TX-18	29	71	35	28	100
Lubbock and rural West Texas: TX-19	71	29	5	20	100
Downtown San Antonio: TX-20	37	63	5	51	100
Parts of Austin and San Antonio connected by Texas H...: TX-21	73	27	3	19	100
Southern Houston suburbs: TX-22	54	46	12	18	100
El Paso to San Antonio: TX-23	38	62	3	51	100
Suburbs north of Dallas and Fort Worth: TX-24	66	34	9	14	100
Part of Austin and rural areas north toward Dallas-F...: TX-25	81	19	5	10	100
Denton County: TX-26	77	23	6	10	100
Corpus Christi and Victoria: TX-27	59	41	5	32	100
Laredo area: TX-28	32	68	5	58	100
Heavily Latino areas north and east of Downtown Houston: TX-29	22	78	13	55	100
Downtown and South Dallas: TX-30	28	72	43	25	100
Williamson and Bell Counties: TX-31	72	28	9	14	100
Suburban North Dallas: TX-32	67	33	10	15	100
Connects Downtown Fort Worth and western Dallas: TX-33	31	69	18	43	100
Brownsville in the Rio Grande Valley and rural count...: TX-34	29	71	2	64	100
San Antonio and heavily Latino areas of Austin: TX-35	44	56	9	40	100
Southern part of East Texas: TX-36	77	23	7	12	100
Ogden and northern Utah: UT-01	90	10	1	5	100
Salt Lake City and southwestern Utah: UT-02	88	12	1	6	100
Provo area and southeastern Utah: UT-03	91	9	0	5	100
Southern Salt Lake County and rural areas to the south: UT-04	86	14	1	7	100
Western Chesapeake Bay and exurbs of D.C. and Richmond: VA-01	73	27	15	6	100
Virginia Beach: VA-02	74	26	15	4	100
Hampton Roads: VA-03	50	50	43	3	100

Table S.9 – Composition of 2016 Electorate, by Race by CD (*continued*)

	White	Non-Whites			All
		All	Black	Hisp.	
Richmond and Southside Virginia: VA-04	56	44	38	3	100
Charlottesville and South-central Virginia: VA-05	81	19	15	2	100
Shenandoah Valley: VA-06	89	11	7	2	100
Suburban Richmond: VA-07	74	26	16	4	100
Inner D.C. suburbs: VA-08	59	41	11	15	100
Southwest Virginia: VA-09	94	6	4	1	100
Loudoun County and D.C. suburbs and exurbs: VA-10	68	32	5	11	100
D.C. suburbs in Fairfax and Prince William Counties: VA-11	54	46	10	16	100
All of Vermont: VT-01	98	2	0	1	100
Northern Seattle suburbs: WA-01	86	14	1	4	100
Northern Puget Sound: WA-02	86	14	2	5	100
Vancouver and southwestern Washington: WA-03	92	8	1	4	100
Yakima and the Tri-Cities: WA-04	75	25	1	15	100
Spokane area and rural eastern Washington: WA-05	93	7	1	3	100
Olympic Peninsula and Tacoma: WA-06	90	10	2	3	100
Seattle: WA-07	78	22	4	5	100
Eastern Seattle suburbs and two rural counties east ...: WA-08	86	14	2	5	100
Southeastern Seattle and the suburbs to its south an...: WA-09	66	34	8	7	100
Olympia area and Tacoma suburbs: WA-10	85	15	3	5	100
Southeastern Wisconsin: WI-01	89	11	3	6	100
Madison area: WI-02	89	11	3	3	100
Southwestern Wisconsin: WI-03	97	3	1	1	100
Milwaukee: WI-04	55	45	28	10	100
Milwaukee suburbs: WI-05	92	8	1	3	100
Fox River Valley and Sheboygan: WI-06	96	4	1	2	100
Northwestern Wisconsin: WI-07	97	3	0	1	100
Northeastern Wisconsin: WI-08	94	6	1	2	100
Northern West Virginia: WV-01	97	3	2	1	100
Charleston to the Eastern Panhandle: WV-02	95	5	3	1	100
Southern West Virginia: WV-03	96	4	3	1	100
All of Wyoming: WY-01	93	7	0	4	100
National	74	26	10	9	100

Table S.10 – Composition of 2020 Electorate, by Race by CD

	White	Non-Whites			All
		All	Black	Hisp.	
All of Alaska: AK-01	73	27	3	4	100
Greater Mobile: AL-01	72	28	24	2	100
Southeastern Alabama: AL-02	68	32	27	2	100
Eastern Alabama: AL-03	74	26	22	1	100
North-central Alabama: AL-04	90	10	5	2	100
Northern Alabama: AL-05	78	22	15	3	100
Birmingham suburbs: AL-06	77	23	16	3	100
Birmingham and the Black Belt: AL-07	37	63	60	2	100
Northeastern Arkansas: AR-01	86	14	11	1	100
Greater Little Rock: AR-02	77	23	18	2	100
Northwest Arkansas: AR-03	83	17	2	7	100
Southern and western Arkansas: AR-04	82	18	13	2	100
Northeastern Arizona: AZ-01	61	39	2	16	100
Southeastern Arizona and eastern Tucson area: AZ-02	67	33	3	19	100
Southwestern Arizona and western Tucson area: AZ-03	38	62	4	39	100
North-central Arizona and Phoenix exurbs: AZ-04	83	17	1	9	100
Phoenix suburbs of Mesa and Gilbert: AZ-05	75	25	3	11	100
Scottsdale and North Phoenix: AZ-06	78	22	2	11	100
Downtown and West Phoenix: AZ-07	36	64	9	36	100
Western suburbs of Phoenix: AZ-08	73	27	3	13	100
Central Phoenix and eastern suburbs: AZ-09	67	33	5	16	100
Northeastern California: CA-01	87	13	1	7	100
Northern California coast: CA-02	80	20	1	9	100
Southern Sacramento Valley: CA-03	61	39	6	19	100
Northern Sacramento suburbs and Sierra Nevada foothi...: CA-04	82	18	1	8	100
Wine country: CA-05	62	38	5	16	100
Sacramento: CA-06	50	50	11	18	100
Suburban Sacramento County: CA-07	63	37	6	11	100
Northern San Bernardino County and the High Desert: CA-08	60	40	7	26	100
Stockton area and eastern Contra Costa County: CA-09	47	53	8	24	100
Modesto area: CA-10	57	43	3	28	100
Contra Costa County: CA-11	56	44	7	16	100
San Francisco: CA-12	54	46	5	11	100
Oakland area: CA-13	43	57	16	14	100
San Mateo County: CA-14	45	55	3	16	100
Eastern Alameda County: CA-15	45	55	5	15	100
Fresno area: CA-16	37	63	6	38	100
Northeast San Jose area: CA-17	37	63	3	14	100
Silicon Valley: CA-18	66	34	2	10	100
San Jose: CA-19	35	65	3	26	100
Monterey Bay: CA-20	49	51	2	30	100
Southern Central Valley and part of Bakersfield: CA-21	35	65	5	40	100
Central Valley near Fresno and Tulare: CA-22	51	49	3	32	100
Bakersfield area: CA-23	59	41	5	25	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

		White	Non-Whites			All
			All	Black	Hisp.	
San Luis Obispo and Santa Barbara:	CA-24	68	32	2	21	100
Northern Los Angeles County suburbs:	CA-25	48	52	7	28	100
Ventura County:	CA-26	57	43	2	31	100
San Gabriel Valley:	CA-27	35	65	5	20	100
Burbank and Glendale:	CA-28	67	33	2	15	100
East San Fernando Valley:	CA-29	29	71	4	43	100
West San Fernando Valley:	CA-30	61	39	4	18	100
San Bernardino area:	CA-31	37	63	10	36	100
East San Gabriel Valley:	CA-32	22	78	3	43	100
Coastal West Side of Los Angeles:	CA-33	71	29	3	9	100
Downtown Los Angeles:	CA-34	18	82	6	43	100
Southwest San Bernardino County and Pomona:	CA-35	21	79	6	45	100
Eastern Riverside County:	CA-36	57	43	4	27	100
Culver City area of Los Angeles:	CA-37	32	68	22	26	100
Southeastern Los Angeles County:	CA-38	20	80	3	44	100
San Gabriel Valley and northern Orange County:	CA-39	39	61	2	25	100
East Los Angeles:	CA-40	8	92	6	60	100
Riverside and Moreno Valley:	CA-41	31	69	8	38	100
Southwestern Riverside County:	CA-42	49	51	5	27	100
Inland cities south of Los Angeles:	CA-43	21	79	21	33	100
South Los Angeles, including Compton:	CA-44	11	89	16	48	100
Irvine area of Orange County:	CA-45	61	39	2	13	100
Santa Ana and the Anaheim area of Orange County:	CA-46	30	70	2	41	100
Long Beach area:	CA-47	44	56	7	23	100
Coastal Orange County:	CA-48	64	36	1	13	100
Southern Orange and northern San Diego Counties:	CA-49	70	30	2	18	100
Inland San Diego County:	CA-50	67	33	2	20	100
South San Diego and Imperial County:	CA-51	23	77	7	52	100
Suburban San Diego:	CA-52	69	31	3	9	100
Eastern San Diego:	CA-53	52	48	7	24	100
Denver:	CO-01	63	37	7	18	100
Boulder and Fort Collins:	CO-02	86	14	1	7	100
Pueblo and the Western Slope:	CO-03	75	25	1	16	100
Eastern Colorado and Denver exurbs:	CO-04	75	25	1	15	100
Colorado Springs:	CO-05	78	22	5	11	100
Aurora and Denver's southern suburbs:	CO-06	66	34	8	12	100
Denver's northern and western suburbs:	CO-07	69	31	1	19	100
Hartford area:	CT-01	69	31	12	9	100
Eastern Connecticut:	CT-02	89	11	3	4	100
New Haven area:	CT-03	76	24	11	7	100
Southern Fairfield County:	CT-04	67	33	10	12	100
Northern Fairfield County and northwestern Connecticut:	CT-05	78	22	6	9	100
All of Delaware:	DE-01	72	28	17	5	100
Pensacola area:	FL-01	78	22	12	3	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Florida panhandle: FL-02	84	16	10	3	100
North Florida, including Gainesville and Ocala: FL-03	77	23	14	5	100
Part of the Jacksonville area and St. Augustine: FL-04	80	20	8	5	100
Jacksonville and Tallahassee areas: FL-05	47	53	45	4	100
Daytona area: FL-06	82	18	8	7	100
Northern Orlando suburbs: FL-07	64	36	9	19	100
Space Coast: FL-08	82	18	8	6	100
Southern suburbs of Orlando: FL-09	48	52	11	32	100
Orlando: FL-10	45	55	25	20	100
North-central Florida and The Villages: FL-11	87	13	5	6	100
Northern Tampa suburbs: FL-12	84	16	4	8	100
St. Petersburg area: FL-13	82	18	9	5	100
Tampa area: FL-14	56	44	15	23	100
Lakeland and the exurbs of Tampa and Orlando: FL-15	68	32	12	14	100
Sarasota area: FL-16	81	19	6	9	100
South-central Florida: FL-17	86	14	5	7	100
Treasure Coast and Palm Beach area: FL-18	75	25	11	10	100
Cape Coral and Fort Myers: FL-19	83	17	5	10	100
West Palm Beach and Fort Lauderdale areas: FL-20	27	73	50	19	100
Southeast Palm Beach County: FL-21	70	30	11	15	100
Boca Raton and northeast Broward County: FL-22	69	31	11	14	100
Broward County-centered: FL-23	49	51	12	32	100
Northern Miami: FL-24	16	84	46	33	100
Northwestern Miami suburbs and eastern Collier Count...: FL-25	29	71	5	63	100
Southwestern Miami area and the Florida Keys: FL-26	21	79	10	65	100
Southern Miami area and Coral Gables: FL-27	26	74	5	64	100
Coastal Georgia: GA-01	69	31	26	3	100
Southwestern Georgia: GA-02	50	50	46	2	100
Southwestern Atlanta exurbs: GA-03	72	28	22	4	100
Eastern Atlanta suburbs: GA-04	28	72	62	5	100
Downtown Atlanta: GA-05	34	66	59	3	100
Northern Atlanta suburbs: GA-06	65	35	14	9	100
Northeastern Atlanta suburbs: GA-07	51	49	18	14	100
South-central Georgia: GA-08	69	31	27	3	100
Northeastern Georgia: GA-09	84	16	5	9	100
East-Central Georgia: GA-10	72	28	23	3	100
Northwestern Atlanta suburbs and exurbs: GA-11	73	27	14	8	100
Augusta and southeastern Georgia: GA-12	64	36	31	3	100
Southwestern Atlanta suburbs: GA-13	29	71	61	5	100
Northwestern Georgia: GA-14	86	14	7	6	100
Honolulu: HI-01	28	72	3	8	100
Northern Oahu and all other islands: HI-02	48	52	2	9	100
Northeastern Iowa: IA-01	95	5	2	2	100
Southeastern Iowa: IA-02	92	8	3	3	100
Des Moines and southwestern Iowa: IA-03	89	11	2	4	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

		White	Non-Whites			All
			All	Black	Hisp.	
	Northwestern Iowa: IA-04	93	7	1	3	100
	Idaho Panhandle and western Boise area: ID-01	87	13	0	6	100
	Boise and Eastern Idaho: ID-02	86	14	0	8	100
	South Side Chicago and southern suburbs: IL-01	44	56	47	6	100
	South Side Chicago and southwestern suburbs: IL-02	39	61	50	8	100
	Southwestern Chicago area: IL-03	72	28	5	16	100
	Heavily Latino areas of Chicago's West and South sid...: IL-04	35	65	4	43	100
	North Side Chicago: IL-05	79	21	3	11	100
	Western Chicago suburbs: IL-06	83	17	2	6	100
	West Side and Downtown Chicago: IL-07	41	59	43	9	100
	Northwestern Chicago suburbs: IL-08	72	28	4	13	100
	Northern Chicago Suburbs: IL-09	76	24	8	7	100
	Chicago's North Shore suburbs: IL-10	74	26	6	11	100
	Southwestern Chicago suburbs: IL-11	67	33	9	15	100
	Southwestern Illinois: IL-12	86	14	11	2	100
	Central Illinois: IL-13	89	11	7	2	100
	Western Chicago exurbs: IL-14	86	14	2	7	100
	East-central and southeastern Illinois: IL-15	95	5	2	1	100
	North-central Illinois: IL-16	92	8	2	4	100
	Northwestern Illinois: IL-17	87	13	7	4	100
	West-central Illinois: IL-18	95	5	2	1	100
	Northwestern Indiana: IN-01	73	27	13	8	100
	North-central Indiana: IN-02	90	10	5	3	100
	Northeastern Indiana: IN-03	91	9	4	2	100
	West-central Indiana: IN-04	93	7	3	2	100
	Northern Indianapolis suburbs: IN-05	86	14	7	2	100
	Southeastern Indiana: IN-06	95	5	2	1	100
	Indianapolis: IN-07	74	26	18	4	100
	Southwestern Indiana: IN-08	95	5	3	1	100
	South-central Indiana and southern Indianapolis suburbs: IN-09	94	6	2	1	100
	Western Kansas: KS-01	89	11	2	6	100
	Topeka, Lawrence, and eastern rural Kansas: KS-02	91	9	3	3	100
	Suburban Kansas City area: KS-03	76	24	8	7	100
	Wichita area: KS-04	85	15	5	5	100
	Western Kentucky: KY-01	93	7	4	1	100
	West-central Kentucky: KY-02	92	8	4	1	100
	Louisville: KY-03	74	26	19	3	100
	Cincinnati suburbs and northern Kentucky border area: KY-04	93	7	2	2	100
	Appalachian Kentucky: KY-05	98	2	1	0	100
	Greater Lexington area: KY-06	88	12	7	2	100
	Suburban New Orleans: LA-01	76	24	11	6	100
	New Orleans and Baton Rouge: LA-02	33	67	58	5	100
	Southwestern Louisiana: LA-03	74	26	21	2	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Northwestern Louisiana: LA-04	65	35	30	2	100
Northeastern Louisiana: LA-05	66	34	31	1	100
Baton Rouge suburbs: LA-06	72	28	21	3	100
Springfield area and the Berkshires: MA-01	83	17	4	8	100
Worcester area and upper Pioneer Valley: MA-02	86	14	4	5	100
Lowell area and the Merrimack Valley: MA-03	73	27	3	13	100
Western Boston suburbs and Bristol County: MA-04	87	13	2	2	100
Boston's northern and outer western suburbs: MA-05	78	22	4	6	100
North Shore: MA-06	85	15	3	6	100
Central Boston area: MA-07	59	41	16	12	100
Southern Boston suburbs: MA-08	79	21	9	3	100
Cape Cod and South Shore: MA-09	91	9	2	2	100
Eastern Shore and Baltimore exurbs: MD-01	85	15	10	2	100
Suburban Baltimore area: MD-02	64	36	26	3	100
Baltimore suburbs to Annapolis to D.C. suburbs: MD-03	69	31	18	5	100
Prince George's County D.C. suburbs: MD-04	29	71	49	11	100
Southern Maryland and outer Prince George's County: MD-05	48	52	38	5	100
Western Maryland and Montgomery County D.C. suburbs: MD-06	69	31	11	7	100
Baltimore and western suburbs: MD-07	45	55	45	3	100
Montgomery County: MD-08	65	35	11	10	100
Portland and Augusta: ME-01	96	4	1	1	100
Northern Maine: ME-02	97	3	0	1	100
Upper Peninsula and northern Lower Peninsula: MI-01	95	5	1	1	100
Western Michigan coast: MI-02	84	16	5	6	100
Grand Rapids area: MI-03	86	14	5	5	100
Central Michigan: MI-04	95	5	1	2	100
Flint, Saginaw, and Bay City: MI-05	81	19	14	3	100
Southwestern Michigan: MI-06	87	13	6	3	100
South-central Michigan: MI-07	93	7	3	2	100
Lansing and northwestern Detroit exurbs: MI-08	87	13	4	4	100
Northern Detroit suburbs: MI-09	84	16	10	1	100
Northern Macomb County and the "Thumb" region: MI-10	93	7	2	2	100
Northwestern Detroit suburbs: MI-11	82	18	4	2	100
Ann Arbor and Dearborn: MI-12	81	19	8	4	100
Western Detroit: MI-13	45	55	46	5	100
Eastern Detroit to Pontiac: MI-14	39	61	53	2	100
Southern Minnesota: MN-01	91	9	2	4	100
Southern Twin Cities suburbs: MN-02	84	16	4	4	100
Western Twin Cities suburbs: MN-03	81	19	6	3	100
St. Paul: MN-04	75	25	8	4	100
Minneapolis: MN-05	71	29	13	6	100
Northern Twin Cities exurbs: MN-06	93	7	2	2	100
Western Minnesota: MN-07	94	6	1	3	100
Iron Range: MN-08	95	5	1	1	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
St. Louis and northwestern suburbs: MO-01	55	45	39	2	100
Suburban St. Louis: MO-02	90	10	3	1	100
East-central Missouri: MO-03	94	6	2	1	100
West-central Missouri: MO-04	93	7	3	1	100
Kansas City area: MO-05	77	23	15	4	100
Northern Missouri: MO-06	92	8	3	2	100
Southwestern Missouri: MO-07	94	6	1	2	100
Southeastern Missouri: MO-08	95	5	3	1	100
Northeastern Mississippi: MS-01	72	28	25	2	100
Mississippi Delta and Jackson: MS-02	37	63	61	1	100
Southwestern and eastern Mississippi: MS-03	64	36	32	1	100
Gulf Coast: MS-04	78	22	17	2	100
All of Montana: MT-01	92	8	0	2	100
Durham and northeastern North Carolina: NC-01	52	48	40	4	100
Raleigh exurbs: NC-02	69	31	18	7	100
Eastern North Carolina coast: NC-03	75	25	17	3	100
Raleigh, Cary, Chapel Hill: NC-04	58	42	21	7	100
Winston-Salem and northwestern North Carolina: NC-05	80	20	13	4	100
Part of Greensboro and rural Piedmont: NC-06	71	29	19	6	100
Southeastern North Carolina: NC-07	74	26	17	5	100
Charlotte suburbs to Fayetteville: NC-08	67	33	21	5	100
South Charlotte to suburban Fayetteville: NC-09	67	33	17	5	100
Asheville and southwestern Piedmont: NC-10	83	17	10	4	100
Appalachian North Carolina: NC-11	91	9	2	3	100
Charlotte: NC-12	44	56	35	10	100
Greensboro and rural western Piedmont: NC-13	72	28	19	4	100
All of North Dakota: ND-01	93	7	2	1	100
Lincoln and rural eastern Nebraska: NE-01	89	11	2	5	100
Greater Omaha area: NE-02	79	21	9	6	100
Western Nebraska: NE-03	90	10	1	6	100
Eastern New Hampshire: NH-01	94	6	1	1	100
Western and northern New Hampshire: NH-02	95	5	1	2	100
Inner Philadelphia suburbs: NJ-01	70	30	13	7	100
Southern New Jersey coast: NJ-02	75	25	9	9	100
Philadelphia suburbs and central Jersey Shore: NJ-03	82	18	9	4	100
Monmouth County area: NJ-04	84	16	4	6	100
Northern Bergen County and northwestern exurbs: NJ-05	74	26	4	10	100
Northern Middlesex County and northern Jersey Shore: NJ-06	63	37	9	13	100
Hunterdon County and New York City's western exurbs: NJ-07	75	25	4	8	100
Jersey City and inner New York City suburbs: NJ-08	42	58	10	27	100
Paterson and inner New York City suburbs: NJ-09	50	50	9	21	100
Newark area: NJ-10	29	71	43	15	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Morris County area: NJ-11	79	21	3	7	100
Trenton and southern Middlesex County: NJ-12	59	41	15	11	100
Albuquerque area: NM-01	47	53	2	33	100
Southern New Mexico: NM-02	51	49	2	34	100
Northern New Mexico: NM-03	45	55	1	31	100
Las Vegas: NV-01	48	52	10	24	100
Northern Nevada: NV-02	70	30	1	15	100
Southern Las Vegas suburbs: NV-03	61	39	7	12	100
Northern Las Vegas suburbs and rural central Nevada: NV-04	52	48	14	18	100
Eastern Suffolk County: NY-01	82	18	4	9	100
Southern Long Island: NY-02	69	31	8	15	100
Northern Long Island: NY-03	76	24	3	6	100
Long Island: NY-04	59	41	14	13	100
Southeastern Queens: NY-05	17	83	51	14	100
Northeastern Queens: NY-06	56	44	6	13	100
Heavily Latino parts of northern Brooklyn and southw....: NY-07	43	57	11	24	100
Eastern Brooklyn: NY-08	37	63	45	11	100
Central Brooklyn: NY-09	42	58	44	7	100
West Side of Manhattan and part of southern Brooklyn: NY-10	78	22	5	8	100
Staten Island and part of southern Brooklyn: NY-11	77	23	6	9	100
East Side of Manhattan and nearby parts of Queens an...: NY-12	78	22	5	8	100
Harlem area: NY-13	20	80	26	33	100
East Bronx and northern Queens: NY-14	35	65	12	30	100
The Bronx: NY-15	4	96	31	40	100
Southern Westchester County and the northern Bronx: NY-16	46	54	30	15	100
Westchester County: NY-17	66	34	9	13	100
Lower Hudson Valley: NY-18	76	24	8	10	100
Hudson Valley: NY-19	91	9	3	3	100
Albany area: NY-20	87	13	7	3	100
North Country: NY-21	95	5	2	1	100
Binghamton and Utica-Rome: NY-22	94	6	2	1	100
Ithaca and southwestern New York: NY-23	95	5	2	1	100
Syracuse area: NY-24	91	9	5	2	100
Greater Rochester: NY-25	79	21	13	4	100
Greater Buffalo: NY-26	81	19	13	3	100
Suburbs and rural areas between Buffalo and Rochester: NY-27	96	4	2	1	100
Suburban Cincinnati: OH-01	77	23	18	1	100
Eastern Cincinnati and east along the Ohio River: OH-02	91	9	6	1	100
Columbus: OH-03	68	32	25	3	100
West-Central Ohio to North-Central Ohio: OH-04	93	7	4	2	100
Northwestern Ohio: OH-05	93	7	2	3	100
Appalachian Ohio: OH-06	97	3	2	1	100
Canton area: OH-07	96	4	3	1	100
Cincinnati suburbs and Springfield: OH-08	92	8	4	1	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

		White	Non-Whites			All
			All	Black	Hisp.	
	Toledo to West Side of Cleveland: OH-09	83	17	10	4	100
	Greater Dayton area: OH-10	84	16	13	1	100
	East Cleveland to Akron: OH-11	50	50	45	2	100
	Northern Columbus area: OH-12	90	10	4	1	100
	Mahoning Valley and Akron: OH-13	90	10	7	1	100
	Northeastern Ohio: OH-14	94	6	3	1	100
	Southern Columbus area and Athens: OH-15	94	6	3	1	100
	Western suburbs of Cleveland and Akron: OH-16	95	5	2	1	100
	Greater Tulsa area: OK-01	82	18	6	5	100
	Eastern Oklahoma: OK-02	87	13	3	3	100
	Northwestern Oklahoma: OK-03	89	11	3	4	100
	Southern Oklahoma: OK-04	87	13	4	4	100
	Oklahoma City area: OK-05	79	21	9	6	100
	Western Portland suburbs and North Coast: OR-01	79	21	1	9	100
	Oregon east of the Cascades and part of southern Oregon: OR-02	88	12	0	7	100
	Eastern Portland area: OR-03	81	19	3	7	100
	Southern Willamette Valley and southern coast: OR-04	91	9	0	4	100
	Salem area and southern Portland suburbs: OR-05	82	18	1	9	100
	Philadelphia's Bucks County suburbs: PA-01	87	13	4	3	100
	Northeast Philadelphia: PA-02	54	46	23	11	100
	West, Northwest, and Center City Philadelphia: PA-03	36	64	57	2	100
	Philadelphia's Montgomery County suburbs: PA-04	81	19	9	2	100
	Philadelphia's Delaware County suburbs and South Phi...: PA-05	69	31	23	2	100
	Philadelphia's Chester County suburbs and Reading: PA-06	78	22	6	9	100
	Lehigh Valley and southern Monroe County: PA-07	81	19	6	8	100
	Scranton/Wilkes-Barre and northeastern Pennsylvania: PA-08	89	11	5	4	100
	Coal Region between Wilkes-Barre and Harrisburg: PA-09	94	6	2	2	100
	Harrisburg and York: PA-10	83	17	9	3	100
	Lancaster and southern York Counties: PA-11	91	9	3	3	100
	Williamsport area and Northern Tier: PA-12	96	4	2	1	100
	South-central Pennsylvania: PA-13	96	4	2	1	100
	Southern Pittsburgh exurbs and southwestern Pennsylv...: PA-14	96	4	3	1	100
	North-central Pennsylvania: PA-15	97	3	2	1	100
	Erie and northwestern Pennsylvania: PA-16	95	5	3	1	100
	Pittsburgh suburbs and Beaver County: PA-17	92	8	5	1	100
	Pittsburgh: PA-18	81	19	15	1	100
	Part of Providence and eastern Rhode Island: RI-01	82	18	5	7	100
	Part of Providence and western Rhode Island: RI-02	87	13	3	6	100
	Charleston and coastal South Carolina: SC-01	76	24	16	4	100
	Augusta and Columbia suburbs: SC-02	72	28	22	3	100
	Northwestern South Carolina: SC-03	82	18	14	2	100
	Greenville-Spartanburg area: SC-04	75	25	16	4	100
	Southern Charlotte suburbs and exurbs: SC-05	72	28	23	2	100
	Black Belt, Columbia, and Charleston: SC-06	46	54	49	2	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
Northeastern South Carolina and the Myrtle Beach area: SC-07	75	25	21	2	100
All of South Dakota: SD-01	92	8	1	1	100
Northeastern Tennessee: TN-01	96	4	2	1	100
Knoxville area: TN-02	91	9	4	2	100
Chattanooga area: TN-03	89	11	7	1	100
Southern part of Middle Tennessee: TN-04	87	13	7	3	100
Nashville: TN-05	71	29	21	4	100
Nashville's eastern suburbs and northern Middle Tenn...: TN-06	93	7	3	2	100
Nashville's southern suburbs and western Middle Tenn...: TN-07	86	14	8	3	100
West Tennessee: TN-08	80	20	15	2	100
Memphis: TN-09	36	64	57	3	100
Tyler and East Texas: TX-01	77	23	13	7	100
Northern Houston and part of west Houston: TX-02	60	40	9	19	100
Collin County: TX-03	67	33	8	9	100
Northeastern Texas: TX-04	83	17	8	6	100
Eastern edge of Dallas and rural areas to its east: TX-05	73	27	12	13	100
Arlington and rural areas south of Dallas: TX-06	62	38	17	13	100
West Houston: TX-07	58	42	11	20	100
Suburbs and exurbs north of Houston: TX-08	77	23	7	12	100
Southwest Houston: TX-09	20	80	37	28	100
Austin to western Houston: TX-10	63	37	9	18	100
Midland and San Angelo parts of rural west Texas: TX-11	76	24	3	17	100
Parker and western Tarrant Counties: TX-12	76	24	7	12	100
Texas Panhandle: TX-13	81	19	4	12	100
Galveston area: TX-14	67	33	16	14	100
McAllen in the Rio Grande Valley and rural counties ...: TX-15	27	73	2	61	100
El Paso: TX-16	22	78	4	65	100
College Station, Waco, and a slice of Austin suburbs: TX-17	73	27	10	12	100
Heavily black parts of central Houston: TX-18	29	71	36	27	100
Lubbock and rural West Texas: TX-19	72	28	5	18	100
Downtown San Antonio: TX-20	34	66	5	53	100
Parts of Austin and San Antonio connected by Texas H...: TX-21	70	30	3	21	100
Southern Houston suburbs: TX-22	49	51	13	18	100
El Paso to San Antonio: TX-23	34	66	4	54	100
Suburbs north of Dallas and Fort Worth: TX-24	63	37	10	14	100
Part of Austin and rural areas north toward Dallas-F...: TX-25	78	22	5	10	100
Denton County: TX-26	72	28	7	11	100
Corpus Christi and Victoria: TX-27	59	41	5	32	100
Laredo area: TX-28	29	71	5	61	100
Heavily Latino areas north and east of Downtown Houston: TX-29	23	77	15	52	100
Downtown and South Dallas: TX-30	27	73	44	24	100
Williamson and Bell Counties: TX-31	67	33	10	16	100
Suburban North Dallas: TX-32	65	35	11	15	100
Connects Downtown Fort Worth and western Dallas: TX-33	31	69	19	40	100
Brownsville in the Rio Grande Valley and rural count...: TX-34	28	72	2	65	100

Table S.10 – Composition of 2020 Electorate, by Race by CD *(continued)*

	White	Non-Whites			All
		All	Black	Hisp.	
San Antonio and heavily Latino areas of Austin: TX-35	42	58	10	40	100
Southern part of East Texas: TX-36	78	22	7	12	100
Ogden and northern Utah: UT-01	85	15	1	9	100
Salt Lake City and southwestern Utah: UT-02	81	19	1	9	100
Provo area and southeastern Utah: UT-03	85	15	0	8	100
Southern Salt Lake County and rural areas to the south: UT-04	76	24	1	12	100
Western Chesapeake Bay and exurbs of D.C. and Richmond: VA-01	69	31	15	8	100
Virginia Beach: VA-02	74	26	15	4	100
Hampton Roads: VA-03	51	49	40	4	100
Richmond and Southside Virginia: VA-04	57	43	35	3	100
Charlottesville and South-central Virginia: VA-05	79	21	16	2	100
Shenandoah Valley: VA-06	87	13	8	2	100
Suburban Richmond: VA-07	70	30	17	5	100
Inner D.C. suburbs: VA-08	61	39	14	11	100
Southwest Virginia: VA-09	93	7	4	1	100
Loudoun County and D.C. suburbs and exurbs: VA-10	63	37	6	11	100
D.C. suburbs in Fairfax and Prince William Counties: VA-11	54	46	12	13	100
All of Vermont: VT-01	96	4	1	1	100
Northern Seattle suburbs: WA-01	80	20	1	5	100
Northern Puget Sound: WA-02	81	19	2	6	100
Vancouver and southwestern Washington: WA-03	87	13	1	5	100
Yakima and the Tri-Cities: WA-04	68	32	1	20	100
Spokane area and rural eastern Washington: WA-05	91	9	1	3	100
Olympic Peninsula and Tacoma: WA-06	85	15	2	4	100
Seattle: WA-07	77	23	4	6	100
Eastern Seattle suburbs and two rural counties east ...: WA-08	78	22	2	6	100
Southeastern Seattle and the suburbs to its south an...: WA-09	61	39	8	8	100
Olympia area and Tacoma suburbs: WA-10	79	21	5	7	100
Southeastern Wisconsin: WI-01	86	14	3	6	100
Madison area: WI-02	88	12	3	4	100
Southwestern Wisconsin: WI-03	96	4	1	1	100
Milwaukee: WI-04	55	45	26	10	100
Milwaukee suburbs: WI-05	90	10	2	4	100
Fox River Valley and Sheboygan: WI-06	94	6	1	2	100
Northwestern Wisconsin: WI-07	96	4	0	1	100
Northeastern Wisconsin: WI-08	93	7	1	3	100
Northern West Virginia: WV-01	96	4	2	0	100
Charleston to the Eastern Panhandle: WV-02	94	6	3	1	100
Southern West Virginia: WV-03	96	4	2	0	100
All of Wyoming: WY-01	92	8	1	4	100
National	72	28	10	10	100

Table S.11 – 2016 Composition of Electorate, by Race by State

	White	Non-Whites			All
		All	Black	Hisp.	
New England					
Connecticut	78	22	8	8	100
Maine	97	3	0	0	100
Massachusetts	82	18	5	6	100
New Hampshire	96	4	1	1	100
Rhode Island	86	14	3	6	100
Vermont	98	2	0	1	100
Middle Atlantic					
New Jersey	67	33	10	12	100
New York	68	32	13	10	100
Pennsylvania	84	16	10	3	100
South Atlantic					
Delaware	75	25	16	5	100
Florida	67	33	12	16	100
Georgia	63	37	27	5	100
Maryland	61	39	26	6	100
North Carolina	71	29	20	5	100
South Carolina	73	27	23	2	100
Virginia	70	30	16	6	100
West Virginia	96	4	2	1	100
East South Central					
Alabama	73	27	23	2	100
Kentucky	90	10	6	2	100
Mississippi	64	36	32	2	100
Tennessee	84	16	12	2	100
West South Central					
Arkansas	82	18	12	3	100
Louisiana	66	34	28	4	100
Oklahoma	85	15	4	4	100
Texas	61	39	10	23	100
East North Central					
Illinois	74	26	12	8	100
Indiana	89	11	6	3	100
Michigan	84	16	10	2	100
Ohio	87	13	9	2	100
Wisconsin	89	11	4	3	100
West North Central					
Iowa	92	8	2	3	100
Kansas	88	12	3	4	100
Minnesota	88	12	4	3	100
Missouri	87	13	8	2	100
Nebraska	90	10	3	4	100
North Dakota	95	5	1	1	100
South Dakota	93	7	1	1	100
Mountain					
Arizona	75	25	3	15	100
Colorado	77	23	3	14	100
Idaho	92	8	0	4	100
Montana	95	5	0	1	100
Nevada	68	32	6	16	100
New Mexico	54	46	2	32	100
Utah	89	11	1	6	100
Wyoming	93	7	0	4	100
Pacific					
Alaska	76	24	2	3	100
California	55	45	5	21	100
Hawaii	44	56	2	7	100
Oregon	88	12	1	6	100
Washington	84	16	2	5	100
National	74	26	10	9	100

Table S.12 – 2020 Composition of Electorate, by Race by State

	White	Non-Whites			All
		All	Black	Hisp.	
New England					
Connecticut	76	24	8	8	100
Maine	97	3	1	1	100
Massachusetts	81	19	5	6	100
New Hampshire	95	5	1	1	100
Rhode Island	85	15	4	6	100
Vermont	96	4	1	1	100
Middle Atlantic					
New Jersey	68	32	10	10	100
New York	69	31	12	10	100
Pennsylvania	83	17	10	3	100
South Atlantic					
Delaware	72	28	17	5	100
Florida	65	35	13	17	100
Georgia	60	40	29	6	100
Maryland	60	40	26	6	100
North Carolina	70	30	19	5	100
South Carolina	72	28	22	3	100
Virginia	69	31	16	6	100
West Virginia	95	5	2	1	100
East South Central					
Alabama	71	29	23	2	100
Kentucky	89	11	6	2	100
Mississippi	64	36	33	1	100
Tennessee	82	18	12	2	100
West South Central					
Arkansas	82	18	11	3	100
Louisiana	65	35	28	3	100
Oklahoma	85	15	5	4	100
Texas	59	41	10	23	100
East North Central					
Illinois	74	26	11	8	100
Indiana	88	12	6	3	100
Michigan	82	18	10	3	100
Ohio	87	13	9	2	100
Wisconsin	88	12	4	4	100
West North Central					
Iowa	92	8	2	3	100
Kansas	85	15	5	5	100
Minnesota	85	15	5	3	100
Missouri	87	13	8	2	100
Nebraska	86	14	4	6	100
North Dakota	93	7	2	1	100
South Dakota	92	8	1	1	100
Mountain					
Arizona	67	33	3	17	100
Colorado	73	27	3	14	100
Idaho	87	13	0	7	100
Montana	92	8	0	2	100
Nevada	60	40	7	16	100
New Mexico	48	52	2	33	100
Utah	81	19	1	10	100
Wyoming	92	8	1	4	100
Pacific					
Alaska	73	27	3	4	100
California	51	49	5	23	100
Hawaii	38	62	2	8	100
Oregon	84	16	1	7	100
Washington	79	21	3	7	100
National	72	28	10	10	100