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 suppplemental 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 × 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 multilingual 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)
- Hillary Clinton (Democrat)
- Gary Johnson (Libertarian)
- Jill Stein (Green)
- Evan McMullin (Independent)

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
- Black or African-American
- Hispanic or Latino
- Asian or Asian-American
- Native American
- Middle Eastern
- Mixed Race
- 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_O:cd_pct_O +
    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 `O` 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).

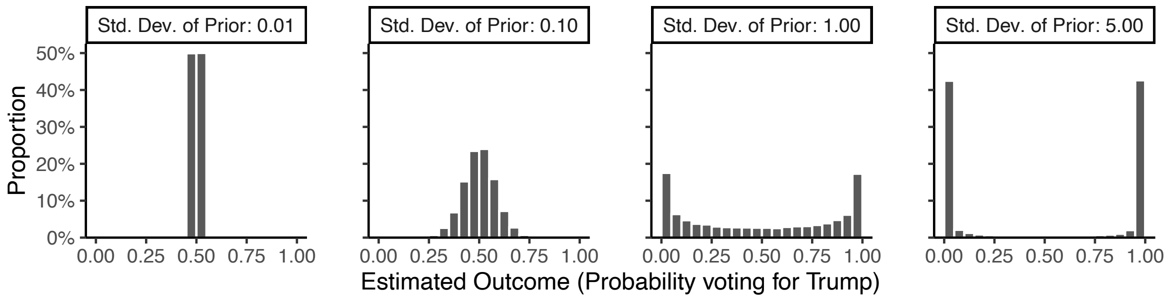
Prior Specification Counter to the intuition that less informative priors imply more flexible values, Figure A.1 shows that less informative priors actually imply more opinionated predicted outcomes (Gelman et al. 2020). To obtain a prior predictive distribution, we sampled only from the prior with the 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

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

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 $\widehat{\tau}_{Wj}$ similarly. Under the standard assumptions, the direct estimator is unbiased and approximately normal $\widehat{\tau}_{gj} \;\dot{\sim}\; \mathcal{N}(0, v_g^2/n_{gj})$. 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

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

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

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

where $w_{gj} = \sigma_g^{-2}/(\sigma_g^{-2} + v_{gj}^{-2} n_{gj})$ 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,

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

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,

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

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 $n_{gj} \longrightarrow 0$, the SAE approach allows the estimate to reduce to

$$\begin{aligned} \hat{\tau}_{Wj}^{\text{SA}} &\longrightarrow \hat{\gamma}_W + \hat{\delta}_W X_j \\ \hat{\tau}_{Bj}^{\text{SA}} &\longrightarrow \hat{\gamma}_B + \hat{\delta}_B X_j. \end{aligned} \tag{9}$$

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

$h \in \{1, \ldots, H\}$. 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 $\widehat{\tau}^{\text{EI}}_{Wj}$ and $\widehat{\tau}^{\text{EI}}_{Bj}$ 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 $\widehat{\tau}^{\text{EI}}_{gj}$ with a least squares regression:

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

and we obtain

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

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 × sex × education × CD] distribution uses 2016 1-year ACS estimates. The [race × 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 × 435]) table into a [160 × 435 × 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 $p(Z \mid \mathbf{X})$. Because $p(\mathbf{X}, Z) = p(Z \mid \mathbf{X})p(\mathbf{X})$, 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 $Z \perp\!\!\!\perp S \mid \mathbf{X}$, we would have that

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

Thus, estimating the relationship between Z and $\mathbf{X}$ in the survey data will provide an unbiased estimate of the joint distribution $\widehat{p}(Z, \mathbf{X}) = \widehat{p}(Z \mid \mathbf{X}, S = 1)p(\mathbf{X})$. 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 $p(Z \mid \mathbf{X})$, we propose to find a probability distribution that satisfies the following equality constraint:

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

which implies that marginally the predicted distribution $p_\beta(Z \mid \mathbf{X}, S = 1)$ 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 $p(Z \mid \mathbf{X}) = p(Z \mid \mathbf{X}, S = 1)$. However, even when the conditional independence assumption does not hold, the above constraint incorporates the population information.

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

Data generating process We consider a case where Z is categorical, so that $p(Z \mid \mathbf{X}, S = 1)$ 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

$$\Pr(Z_i = k \mid \mathbf{X}_i, S_i = 1) = \frac{\exp(\boldsymbol{\beta}_k^\top \widetilde{\mathbf{X}}_i)}{\sum_{k'=1}^{K} \exp(\boldsymbol{\beta}_{k'}^\top \widetilde{\mathbf{X}}_i)}$$

where we set $\boldsymbol{\beta}_1 = \mathbf{0}$ for identification. Here, $\widetilde{\mathbf{X}}_i$ includes the intercept as well as interaction terms between variables.

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

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

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

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

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

where the product is over respondents in the survey data.

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

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

where

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

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 $\boldsymbol{\beta}_{-k}$, and try to estimate $\boldsymbol{\beta}_k$ (coefficients for the kth category). By the EM algorithm derived in Yamauchi (2021), we have the following M-step,

$$\begin{aligned} \text{maximize} \quad & Q_k(\boldsymbol{\beta}_k) \\ \text{subject to} \quad & g_k(\boldsymbol{\beta}_k) = 0 \end{aligned}$$

where the objective function is quadratic in $\boldsymbol{\beta}_k$ after conditioning on the augmented Polya-Gamma random variable ω.

$$Q_k(\boldsymbol{\beta}_k) = -\frac{1}{2}\boldsymbol{\beta}_k^\top(\mathbf{S} + \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\beta}_k + \boldsymbol{\beta}_k^\top(\widetilde{\mathbf{X}}^\top \boldsymbol{d} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0)$$

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

We can obtain the update by considering the Lagrangian

$$L(\boldsymbol{\beta}_k, \lambda) = Q_k(\boldsymbol{\beta}_k) + \lambda g_k(\boldsymbol{\beta}_k).$$

The optimality conditions require we solve for

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

where each component of the gradient is

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\beta}_k} L(\boldsymbol{\beta}_k, \lambda) &= -(\mathbf{S} + \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\beta}_k + (\widetilde{\mathbf{X}}^\top \boldsymbol{d} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0) + \lambda \frac{\partial}{\partial \boldsymbol{\beta}_k} g_k(\boldsymbol{\beta}_k), \\ \frac{\partial}{\partial \lambda} L(\boldsymbol{\beta}_k, \lambda) &= Q_k(\boldsymbol{\beta}_k) + g_k(\boldsymbol{\beta}_k). \end{aligned}$$

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

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

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

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

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

Relaxing the exact constraint The exact constraint $g(\boldsymbol{\beta}) = \mathbf{0}$ 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

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

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 $\boldsymbol{\beta}$, 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 $\widehat{\boldsymbol{\beta}}$ for the conditional distribution, we obtain the population joint probability as

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

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

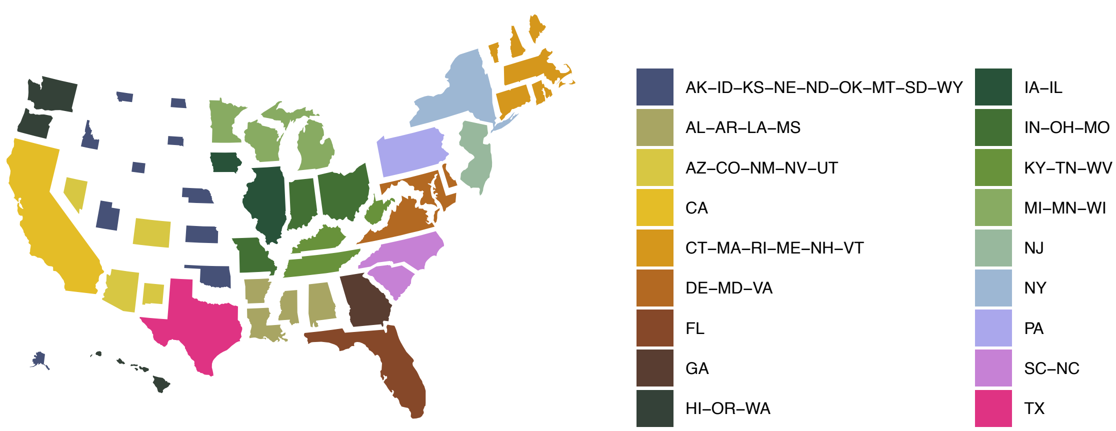
We start with the 4-way demographic table from above and wish to estimate the turnout rate in each covariate set s. We again use the CCES, where turnout is 1 if the voter matches to the Catalist voter file in the state and 0 otherwise. This validated vote variable is a standard variable in the public 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.

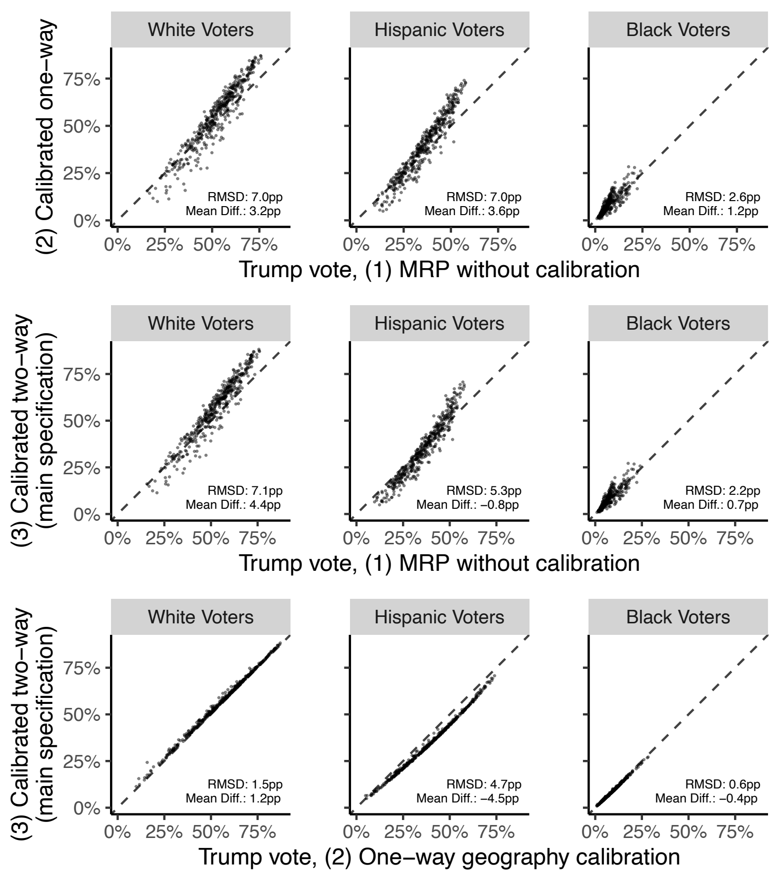
The hierarchical modeling that is used in the main CCES specification can in theory be used to overcome such small sample problems. The reason why we group states in this preparation stage only is because the outcome balancing multinomial model does not partially pool, and limited by the interactions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

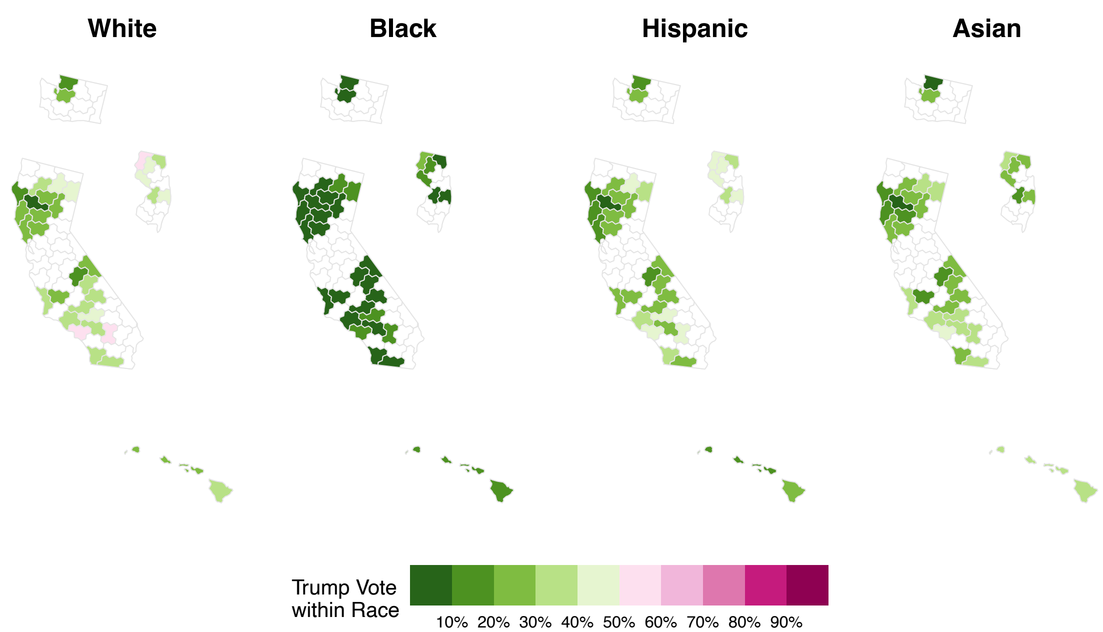


Table B.1 – Comparison of Survey Estimates. All estimates show two party Republican vote in the Presidential election. Exit Polls use the National Election Poll (as published in CNN). *Catalist* refers to Catalist's MRP estimates. *LD* refers to Latino Decisions and the American Election Eve Poll. *CMPS* 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 Hispanic 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.

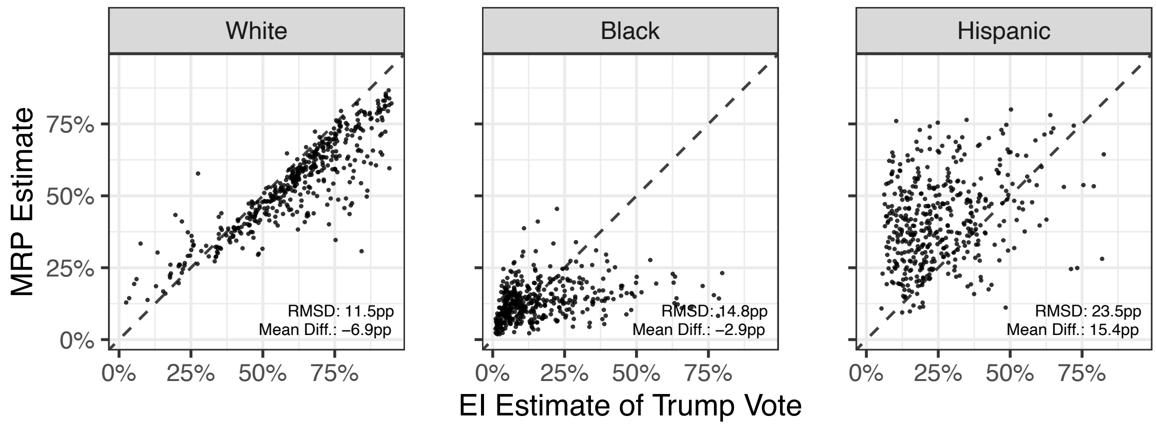
MRP methodology The CCES includes a variable for party registration that comes directly from the Catalist voter file match. In other words, in the CCES party registration is observed without error. We estimate a MRP model predicting party registration in Florida with the same methods described in this 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.

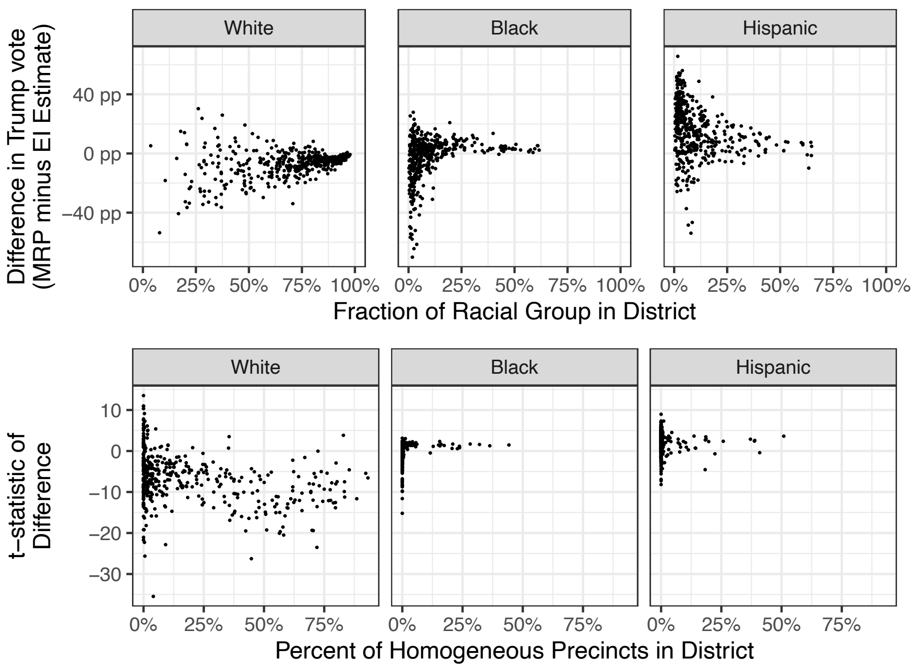


precinct-level which are then assigned to districts as of 2020. We use the combined file from McCartan et al. (2022), which incorporates election data and shapefiles from Voting and Election Science Team (2020) with Census demographics. We compare our method's and EI's estimate for the 2020 election 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.

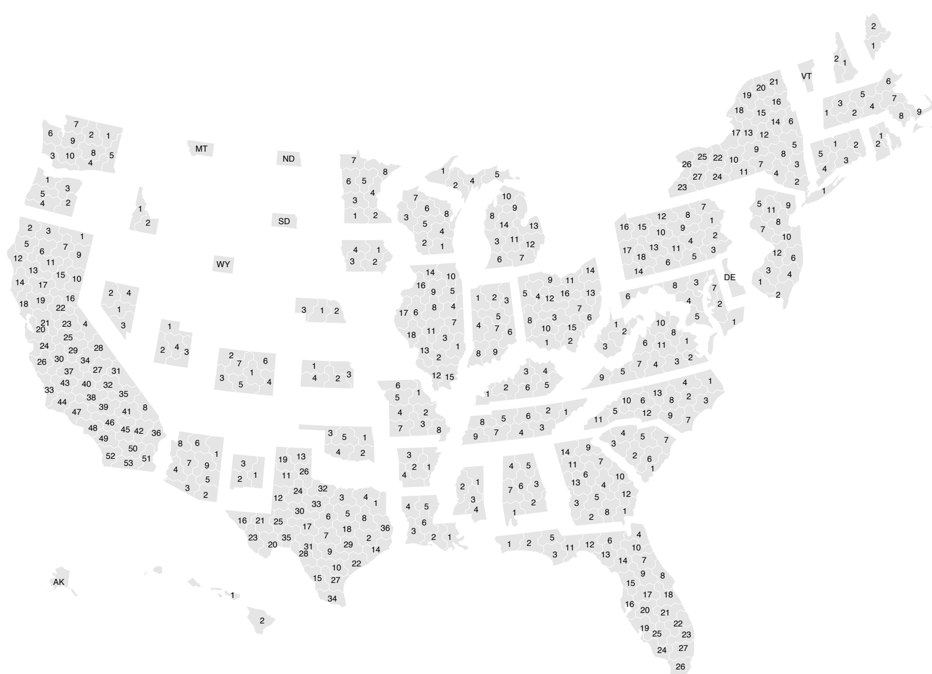
C Additional Estimates and Standard Errors

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/MAZNJ6` (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 inthe 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

	White	Non-Whites All	Non-Whites Black	Non-Whites 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

	White	Non-Whites All	Non-Whites Black	Non-Whites 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

	White	Non-Whites All	Non-Whites Black	Non-Whites 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.

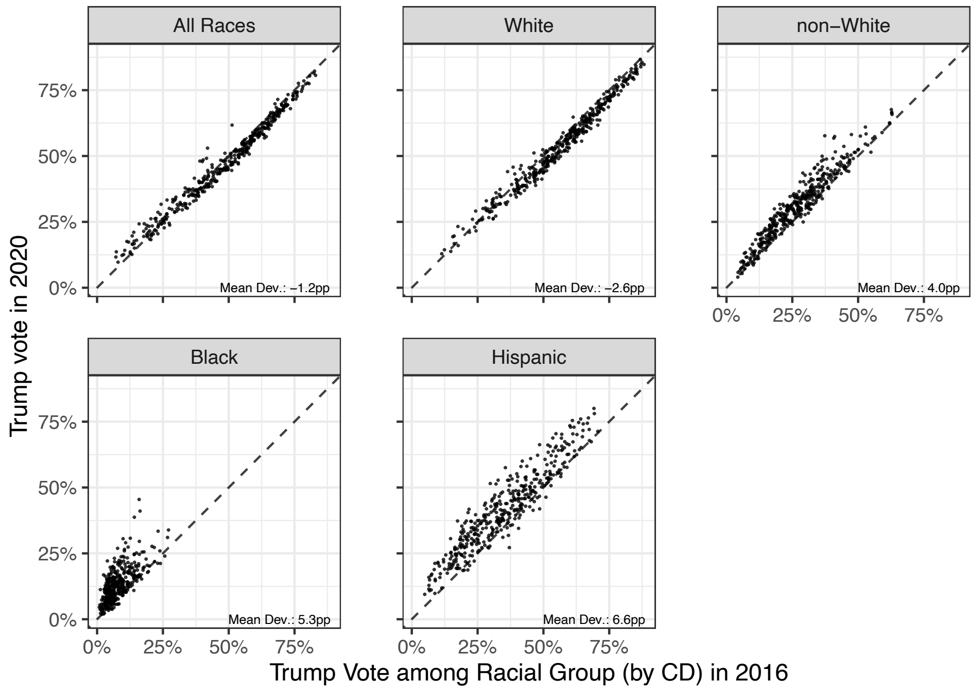


Figure C.3 – Distribution of CD-level Racial Gap Estimates. A histogram of the distribution shown in Figure 3.

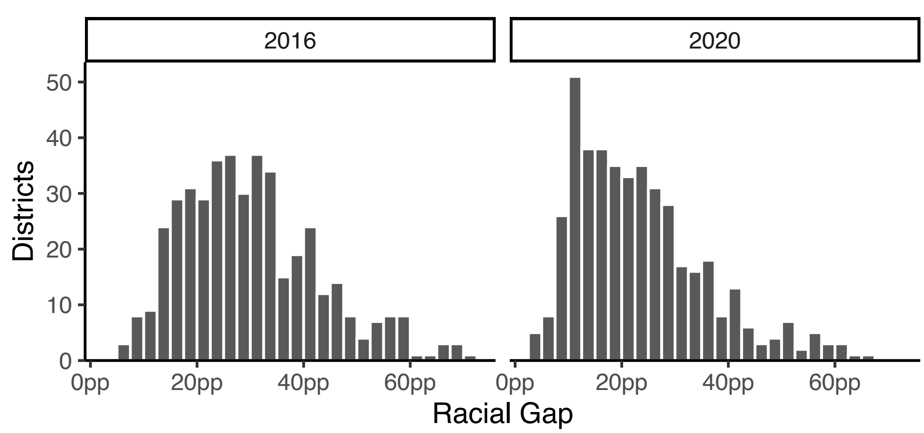
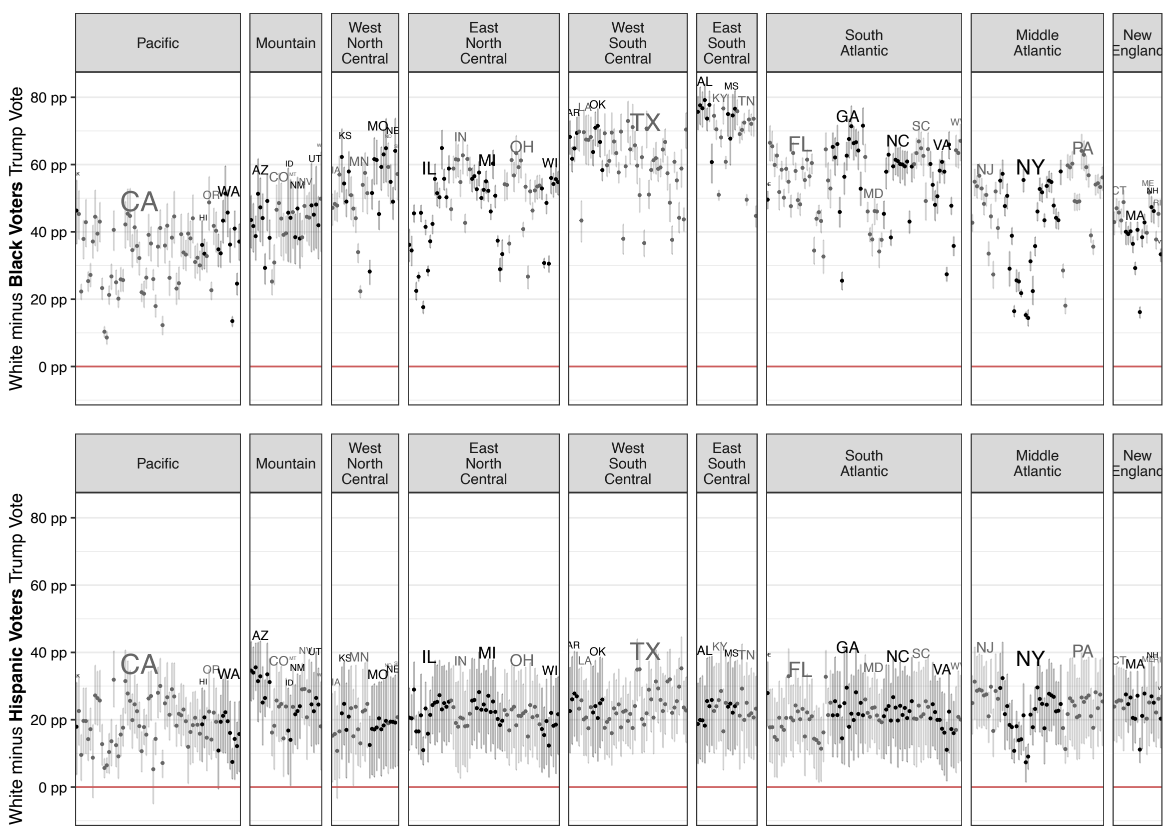


Table C.2 – Summary Statistics of CD-level Racial Gap

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 percent 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

		Non-Whites			
	White	All	Black	Hisp.	All
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

(b) 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

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.