Pre-Analysis Plan:
Predicting Racial Descriptive Representation in the U.S. Congress: The Case of the 2026 Midterm Elections*

Yuki Atsusaka† Kolbe A. Dumas‡

First version: October 27, 2026
This version: November 2, 2026

Abstract

Racial descriptive representation in the U.S. Congress has shaped debates over redistricting, voting rights, and partisan politics. We develop theory-based predictions of racial descriptive representation in the 2026 congressional elections. Drawing on a recent theoretical model with two inputs—district racial composition and the prior electoral performance of co-ethnic candidates—we generate district-level probabilistic predictions of winners' racial group before Election Day. We predict that, in some districts, a focal racial group has a high probability of electoral victory even when it does not constitute a numerical majority. Nationally, we predict that the U.S. House will elect, on average, 7 Asian, 40 Black, 39 Latino, and 349 White legislators. These predictions imply White overrepresentation in the House in the 120th Congress by 14 percentage points, alongside Asian, Black, and Latino underrepresentation by 3, 3, and 5 percentage points, respectively, relative to each group's share of the citizen voting-age population. If accurate, the forecasts will demonstrate the predictive value of local electoral context beyond district racial composition alone.

Remark: This is a pre-analysis plan. Our predictions will be evaluated in public after Election Day. This study is purely academic and intends neither to inform nor influence election results in any way, shape, or form.

*For helpful feedback, we thank Hane Lee. Our district-level predictions are available at https://logical-model.github.io/predictions-2026. We use GPT-5.6 Terra (medium) to assist with coding for our data construction and predictions.

†Hobby School of Public Affairs, University of Houston. https://www.atsusaka.org/

‡Department of Political Science, University of Houston. https://kolbedumas.com/

1 Introduction

Racial descriptive representation in the U.S. Congress has had enormous implications for redistricting, voting rights, and partisan politics (Grofman, Handley and Niemi, 1992; Davidson and Grofman, 1994; Lublin, 1997; Grose, 2011). The politics and debates over legislative redistricting, including the creation of majority minority districts, often center on whether likely winners are from a specific racial group (Cameron, Epstein and O'halloran, 1996; Hicks et al., 2018; Bullock III, 2010; Canon, 2022). The presence of racial minority representatives has also been a cornerstone of the post-Voting Rights Act (VRA) of 1965 legal and political debates (Handley and Grofman, 1994; Guinier, 1994; Thernstrom, 1987; Schuit and Rogowski, 2017; Kenny et al., 2026). Moreover, racial representation has had direct and indirect consequences for partisan balance in Congress and public policy (Epstein and o'Halloran, 1999; Grose, 2011; Schuit and Rogowski, 2017).

Consequently, predicting racial representation at the district, state, and legislature levels has been of practical and theoretical importance in the study of American democracy. However, previous debates have provided us with surprisingly few ways to make such predictions. Nevertheless, most discussions and studies do not go beyond making conjectures about the association between district demographics and representative race. In legal contexts, debates on the "65% rule," "majority-minority districts," "minority-opportunity districts," and "minority-influence districts" appear to equate racial composition with racial representation (Brace et al., 1988; Shaw v. Reno, 1993; Georgia v. Ashcroft, 2003; League of United Latin American Citizens v. Perry, 2006). In social science, while many studies examine the explanatory power and/or causal effects of multiple variables on the presence of Black, Latino, and Asian officeholders, their goal is not to make probabilistic district-level predictions about minority representation (Lublin, 1997; Casellas, 2011; Lublin and Wright, 2024; Juenke and Shah, 2016; Fraga, Juenke and Shah, 2020; Shah, 2014; Lee and Velez, 2024; Abott and Magazinnik, 2020). Still others develop and make theoretical predictions of minority representation (Grofman, Handley and Lublin, 2001; Lublin et al., 2019; Atsusaka, 2021; Hankinson, Loffredo and Magazinnik, 2026), but their applications are often limited to specific cases and explanatory.

In this work, we seek to make theory-based predictions of racial descriptive representation in the 2026 congressional elections. Drawing from a recent theoretical model of strategic candidate entry with only two terms (Atsusaka, 2021), we make probabilistic predictions about the racial identification of winners *before* Election Day and evaluate the accuracy of such predictions after official election results are released. This theoretical model assumes

that candidates are short-term instrumental such that their sole goal is to win seats in the upcoming elections, and they run for office by accounting for how likely they are to win seats. The theory then states that such rational candidates estimate their odds of winning by using the two pieces of information from (1) district racial composition and (2) electoral performance of co-ethnic candidates in the previous elections. Practically, the model allows us to make *theoretical* and *probabilistic* predictions (unlike parametric or classification models) about the presence and number of officeholders from different racial groups in a specific district, state, and legislature. While previous applications of the model focus on local and state legislative elections (Atsusaka, 2021; Hankinson, Loffredo and Magazinnik, 2026), we apply the model of strategic candidate entry to the U.S. House of Representatives elections.

Leveraging the theoretical model with just two predictors, we make district-level probabilistic predictions about the racial identification of winners before Election Day. We predict some districts to have high probabilities of minority electoral victory even when the target group does not constitute a numerical majority. We highlight that we do not predict any group's "candidates of choice." Similarly, we do not assume that co-ethnic candidates are voters' preferred candidates. Politics is multidimensional, and voters are most likely to account for many other factors beyond candidate race.

Aggregating our district-level results, we predict that the highest proportion of minority representatives will be in Georgia (5), New York (5), Illinois (3), and Texas (3) (African Americans), California (17), Texas (approximately 8), Florida (3) (Latinos), and California (5) and New York (1) (Asian Americans), respectively. Nationally, we predict the election of 7 Asian, 40 Black, 39 Latino, and 349 White legislators in the U.S. House, forecasting sizable White overrepresentation in the 120th Congress (by 15 percentage points) along with moderate Asian, Black, and Latino underrepresentation (by about 3, 3, and 5 percentage points) relative to their group shares in the citizen voting-age population (CVAP).

After Election Day, we will evaluate the predictive accuracy of our model with several criteria. To "preregister" our evaluation, we assess our predictions based on *hypothetical ground-truth* data, where the winner's race coincides with the plurality group of a district. At the district level, we leverage multiple standard metrics for evaluating probabilistic predictions and classifications (Kuhn, Johnson and Vaughan, 2013). For our state- and national-level predictions, we leverage graphical evaluations and several standard metrics.

If accurate, our work has several implications. Most importantly, while existing studies and debates primarily focus on a set of demographic thresholds that may determine racial

descriptive representation in a district, our results suggest that district racial composition is only half of the story. We find that the racial margin of victory—the relative performance of co-ethnic candidates in previous elections—provides key information about how likely winners of different races would emerge in each district. Consistent with previous findings (Atsusaka, 2021), we illustrate that the percentage of voters from a given group required to elect the co-ethnic candidate depends on the context, while also quantitatively measuring such a dependency.

2 Theoretical Model

What is the probability that a specific legislative district will elect a candidate with a particular racial background? To answer this question, we leverage a recently proposed theoretical model of minority representation by Atsusaka (2021). It is a parsimonious mathematical model that is (1) constructed based on a massive theoretical knowledge accumulated in the literature on race and representation and (2) designed to make probabilistic predictions about racial descriptive representation at the district level.[1] We call it the strategic entry model of racial representation based on the theoretical argument the model makes. We caution that the theoretical model was originally developed for single-member districts with plurality rule or first-past-the-post without primaries. Thus, the application of the model to U.S. general congressional elections may produce incorrect predictions due to the difference in underlying mechanisms for candidate nomination and partisan structures. Nonetheless, one of our goals is to examine to what extent such a model may accurately predict congressional races and explain why it does (or does not).

The strategic entry model first posits that district-level descriptive representation can be modeled by the behavior of the most viable candidates from different racial groups. That is, if the most "winnable" candidate from a given group decides not to run for office in a district, that district will not have any candidates from the group. If so, if we understand when such viable candidates enter the contest, it is possible to understand why some districts have, for example, Black candidates and others do not.

With this premise, the model makes several assumptions about under what conditions

[1]This "logical" model is an example of the methodological approach to constructing *quantitatively predictive logical models*, which has been developed primarily in the literature on electoral systems (Taagepera, 2007, 2008; Shugart and Taagepera, 2017). The goal of logical models is to explain political phenomena by using logic, theory, and substantive knowledge about the topic of interest before (blindly) relying on empirical data.

candidates of different backgrounds will enter the race (Atsusaka, 2021, 1212-13). The central idea is quite simple:

> The most viable candidate from a specific group is more likely to run for office when the perceived chance of winning is higher.

The model then argues that such candidates estimate their chance of winning with two pieces of information: (1) district racial composition and (2) electoral performance of co-ethnic candidates in the same district (or the racial margin of victory) from the immediate last election. Atsusaka (2021) then theorizes that these candidates may consider two extreme scenarios: (a) the highest level of racially polarized voting occurs and therefore district racial composition determines the winner and (b) the upcoming election will repeat its previous result. Both scenarios serve as benchmark cases in the theory. The theory then makes the following second argument:

> The most viable candidate from a specific group estimates the chance of winning by taking the average of the two extreme scenarios.

More formally, Atsusaka (2021) introduces the following logical model:

$$\Pr(\text{Target Group Winning}) = \Phi(\sqrt{M_{t-1}^{\text{adj}} C_t} - 50), \tag{1}$$

where C is the percentage of voters from a target racial group (e.g., % Blacks), M^{adj} is the adjusted racial margin of victory for the target group (the racial margin + 50)—quantifying how close the group is to securing its descriptive representation, and Φ is the cumulative distribution function (CDF) of the standard normal distribution. The index denotes whether the quantity is defined at time t (e.g., year 2026) or at time $t-1$ (e.g., year 2024). The "-50" part is simply for centering the quantity around zero. The square root represents the idea of taking the (geometric) average of the two theoretical bounds.

In our application, the two terms are:

- M_{t-1}^{adj}: the adjusted racial margin of victory for the target group in 2024
- C_t: % of voters from the target group in 2026

Finally, the adjusted racial margin of victory is defined as follows:

$$M^{\text{adj}} = M^{\text{raw}} + 50 = \frac{1}{2}\left(v^{\text{target}} - v^{\text{non-target}}\right) + 50, \tag{2}$$

where v^{target} and $v^{\text{non-target}}$ are the vote shares of the highest-vote-share (most viable) candidates from the target and its comparison groups, respectively. More technically, the racial margin of victory is the minimum share of voters who need to vote differently to change the race of the winner, and it *summarizes* numerous factors that prior studies of minority representation hypothesize to affect minority electoral victory (Atsusaka, 2021, 1213-16). In other words, M summarizes all the local political factors that may affect racial representation beyond district demographics.

3 Data

We compile four datasets of 435 U.S. House districts that we use for predicting racial descriptive representation for African Americans, Asians, Latinos, and Whites, respectively. Our dataset includes two key pieces of information: (1) the district racial composition as of 2026 and (2) the racial margin of victory in the 2024 congressional elections.

First, we use 2020–2024 American Community Survey CVAP estimates tabulated to the district boundaries in effect for the 2026 election. The four groups are mutually exclusive: African Americans, Asians, and Whites are all non-Hispanic and Asians include Pacific Islanders. As we will discuss below, we account for mid-decade redistricting and match the CVAP data with the latest district boundaries that are used for the 2026 midterm elections. The states that experienced mid-decade redistricting include Alabama, California, Florida, Louisiana, North Carolina, Ohio, Tennessee, Texas, and Utah.

Second, we assembled data on 1,128 named candidates across all 435 U.S. House districts in the 2024 election to obtain the racial margin of victory. Candidate vote totals and election outcomes come from the House Clerk's official election statistics. We collected racial and ethnic descriptions from multiple sources, including Rutgers' Center for American Women and Politics, FiveThirtyEight's candidate research, the House Historian, NALEO, academic replication archives, official biographies, candidate statements, and news reporting. These sources provide 915 attributed records covering 674 candidates (59.8%), leaving 454 without published-source descriptions. We retained the original descriptions, supporting passages, and attribution, linking records through documented names, offices, and districts. Eight additional winners received manual visual classifications through a portrait review dashboard, with their coding basis recorded separately. Published descriptions and manual assessments therefore cover 682 candidates, including all 435 winners.

For candidates lacking published descriptions, we collected portraits through a prede-

fined agent search protocol specifying eligible publishers, search order, candidate matching requirements, and image quality criteria. Candidate names and office information establish the documentary link between each person and photograph. Original images, source pages, search records, and file hashes are archived so reviewers can verify the collected materials and repeat the search procedure.

We integrate these photographic data with the recently developed image-based prediction methods (Lee and Velez, 2025) to code the remaining candidates' race. A DeepFace implementation adapted from the PredEthnorace repository generated racial and ethnic category scores for 347 candidates, including the eight manually coded winners.[2] The implementation uses candidate portraits, an adaptation of the published study's procedure for collecting faces associated with names. We will assign each candidate a model classification using the category with the highest score, while retaining all six scores in the dataset. Model classifications will remain separately identified from published-source descriptions and manual assessments throughout the analysis. These outputs add 339 candidates beyond those covered by source descriptions or manual assessments. Combined record coverage reaches 1,021 candidates (90.5%), leaving 107 with neither source descriptions, manual assessments, nor model outputs.

Before constructing the prediction inputs, we will review candidate links and reconcile the available descriptions with the study's racial and ethnic categories. Earlier election records require confirmation that they refer to the same candidate, while photographic records require review of the documented person and image association. Overlapping source identities remain visible, and a White description alone does not establish non-Hispanic White identity. The model's six categories also omit Native American and Pacific Islander classifications.

Figure 1 shows that almost all uncoded candidates are weak candidates receiving less than 5% of the vote. The figure also clarifies that even when their vote shares are leveraged, their contributions are small. In such a situation, we code the remaining candidates as White.

There are several notable caveats and limitations. First, this study primarily focuses on the four racial groups, while not being able to account for other groups, including mixed-race legislators. Second, mid-decade redistricting in several states between 2024 and 2026 raises a serious challenge for making predictions. Our theoretical model assumes the presence of continuity between "two constituencies" in the same district across two contests, and thus does not account for redistricting directly. To address this issue, we construct the racial

[2]We use a Python-based tool available at https://github.com/DianaDaInLee/PredEthnorace.

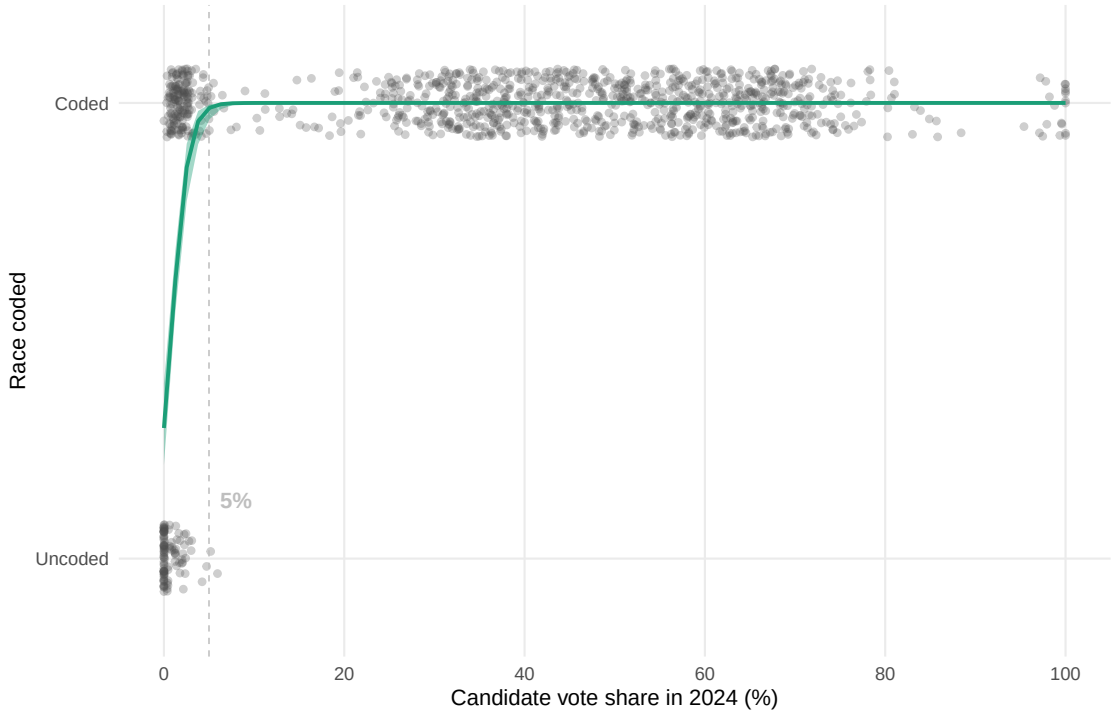


Figure 1: Candidate Vote Share and Racial Coding (with a Logit Curve)

margin of victory in 2024 for the new 2026 districts by reallocating candidate vote shares from the old districts to the new districts. More specifically, each old-district candidate's vote total is allocated across new districts in proportion to the old district's 2020 Census population located within each new district, which we use to compute the racial margin of victory.

4 Predictions

4.1 District-Level Predictions

We begin by making predictions at the district level. For each district, we predict the probability of observing a winner from a given racial group.[3] Figure 2 visualizes our predictions as a function of the two predictors across four racial groups. The upper-right area with redder colors shows where our predicted probability is 0.9 or higher. The bottom-left area with bluer colors is where we predict a probability of 0.1 or lower. In the middle of each panel,

[3]We generate our predictions using the open-source R package `logical` available at https://github.com/YukiAtsusaka/logical.

there is a narrow band where the predicted probability is between 0.1 and 0.9.

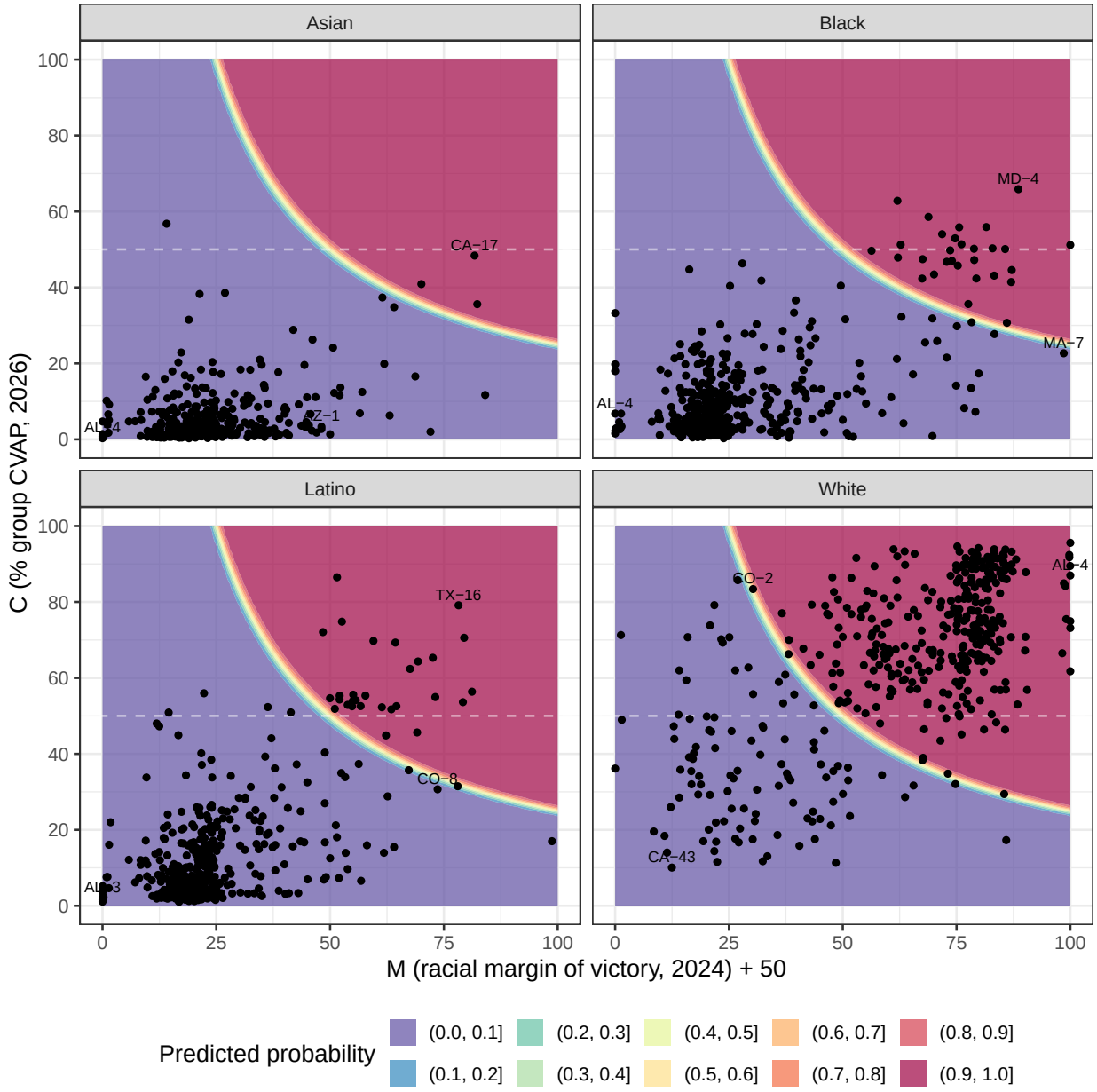


Figure 2: District-Level Predicted Probability of Electing Candidates of Different Races

These predictions are the foundation of our predictive inference. The figure also shows that our probabilistic predictions are much more nuanced than a naive classification based solely on the 50% population threshold (denoted by a horizontal dashed line). Even in minority-minority districts, for example, we predict the victory of Black candidates. In contrast, some districts are predicted to have winners from a given group even when the group does not constitute a numerical majority. In fact, this granularity is the most notable

with White candidate success.

4.2 State-Level Predictions

Next, we aggregate our district-level predictions up to the state level. More specifically, we predict the number of legislators in each state across the four groups. A key challenge is that it is unclear how to leverage our group-versus-rest predictions (e.g., whether a Black candidate is elected) into seat-level predictions (e.g., which racial group wins descriptive representation). Because previous research does not offer practical guidance (Atsusaka, 2021; Hankinson, Loffredo and Magazinnik, 2026), we experiment with a novel approach. We caution that its validity needs to be further examined in future research, and our predictions should be taken with a grain of salt. With the group-specific probabilities, p_{A}, p_{B}, p_{L}, and p_{W} for Asians, Blacks, Latinos, and Whites, respectively, we normalize each probability as follows:

$$\tilde{p}_g = \frac{p_g}{p_{\mathrm{A}} + p_{\mathrm{B}} + p_{\mathrm{L}} + p_{\mathrm{W}}}, \tag{3}$$

where $\tilde{p}_g$ and p_g are the normalized and raw probabilities that a given district elects a candidate from group g. To simulate predicted seats, we then randomly draw from a categorical distribution with a vector of success probabilities for the four groups. Here, it is important to emphasize that we treat the four racial groups as mutually exclusive and collectively exhaustive.

Let R_i be the race of the winner in district i and $\tilde{p}_{\mathrm{A},i}$, $\tilde{p}_{\mathrm{B},i}$, $\tilde{p}_{\mathrm{L},i}$, and $\tilde{p}_{\mathrm{W},i}$ be the four normalized probabilities. Our approach is to repeatedly draw samples from the following categorical distribution.

$$R_i \sim \text{Categorical}(\tilde{p}_{\mathrm{A},i}, \tilde{p}_{\mathrm{B},i}, \tilde{p}_{\mathrm{L},i}, \tilde{p}_{\mathrm{W},i}) \tag{4}$$

In this work, we draw 20000 samples from the above distribution and summarize the results with mean and simulated 95% confidence intervals.

Figure 3 presents our state-level predictions. States are sorted by the number of seats available. The gray polygons represent "forbidden areas" where no value can logically appear (Taagepera, 2008). All small-seat states (with one seat or two) are predicted to have all White representation, except for Hawaii. With the exception of New Mexico, in all other states, the largest number of seats are predicted to be held by White representatives. Our predictions

also show that Asian, Black, and Latino representation seem more likely in states with medium-sized or large House delegations, with New York, Florida, Texas, and California as notable cases. Asian (5 seats) and Latino (17 seats) representation are predicted to be the largest in California regarding the number of seats, whereas Black representation is predicted to be the most prominent in Georgia and New York (5 seats) according to our model.

Compared to the district-level predictions, predicting the number of legislators of each race appears more challenging because the latter depends not only on the accuracy of group-specific predictions, but also on the performance of our predictions for all four groups combined. The goal of this exercise is thus to examine in which states our predictions may be more accurate than in other states and explore the reason why.

4.3 National-Level Predictions

Finally, we make predictions about racial descriptive representation at the national level. We aggregate the number of predicted representatives from the four groups at the state level to the national level (with 20000 random draws). Making predictions at the national level appears to be the most challenging task, as they depend on both the accuracy of our district- and state-level predictions for all four groups. Any systematic prediction failures may drastically generate poor performance, especially when misleading predictions for one group may automatically spill over to another group.

Figure 4 visualizes the geographical distribution of the U.S. congressional districts with the most likely racial representation. The average predicted composition of the U.S. House includes 7 Asian, 40 Black, 39 Latino, and 349 White legislators. As our state-level predictions show, the uncertainty around our predictions is quite small, but the simulation intervals are narrow conditional on the specified model and inputs. Overall, we predict that about 80% of congresspeople are White and about 20% are racial minorities. The figure clarifies that, except for White representatives, racial representation tends to occur in a subset of states, such as California for Asian representation, Southern states and Northern states with urban districts for Black representation, and Southwestern states for Latino representation.

What do our predictions tell us about the degree of racial descriptive representation at the national level? The right panel of Figure 5 compares the share of different racial groups in the CVAP as of 2026 to the share of the same groups in the U.S. House based on our predictions. For a comparison, we also make the same contrast with the 2024 races. We predict that the U.S. House in the 120th Congress is characterized by White overrepresentation by 14

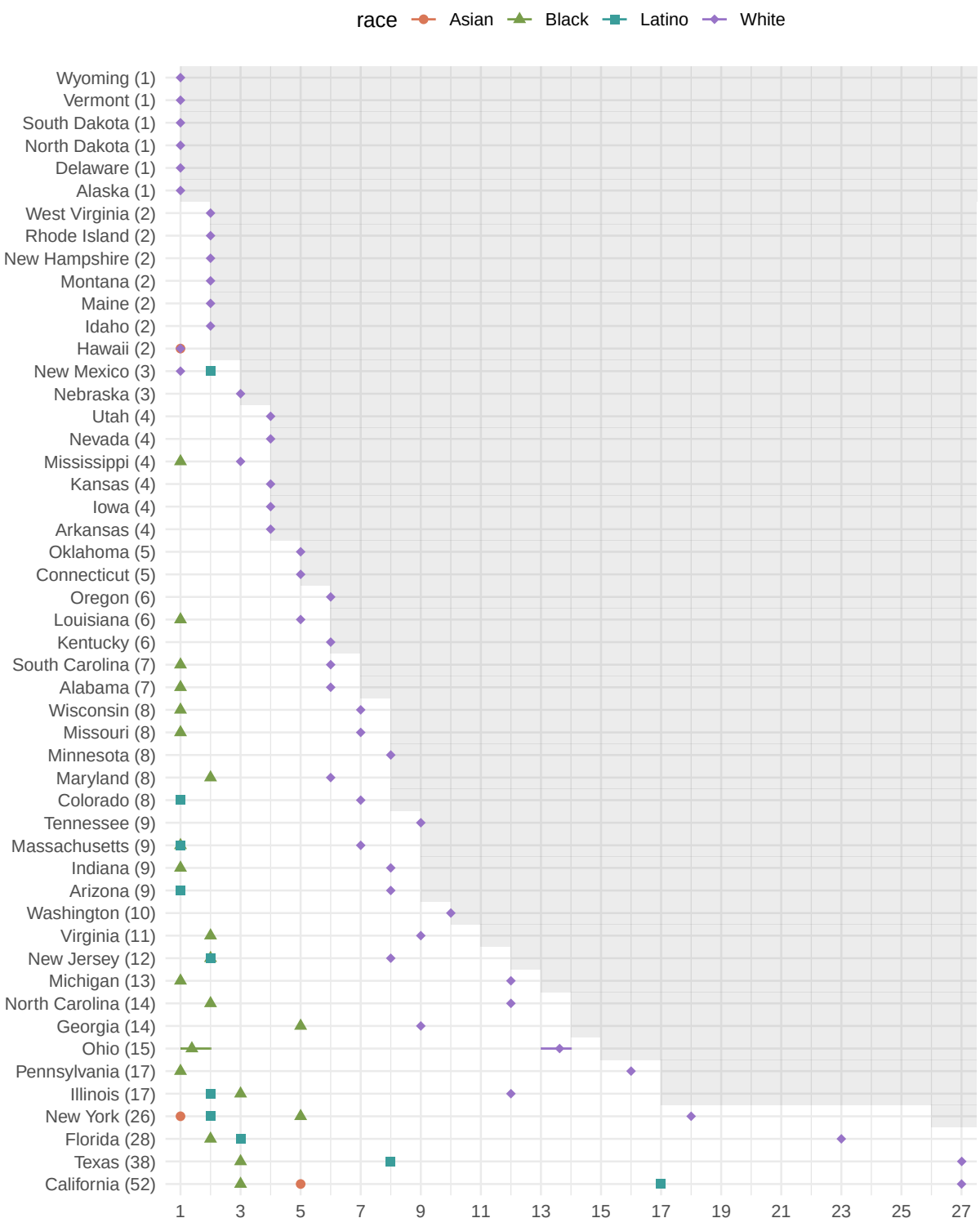


Figure 3: Predicted Number of Legislators by Race and State (Mean and 95% Confidence Intervals)

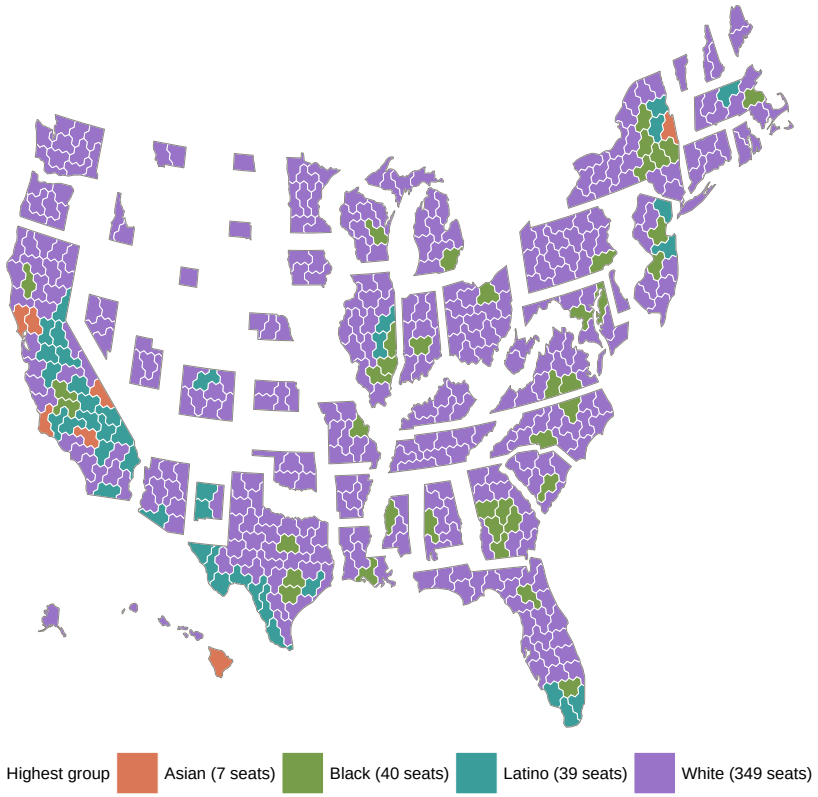


Figure 4: Predicted Geographical Distribution of Racial Representation in the U.S. House

percentage points and Black, Latino, and Asian underrepresentation by 5, 3, and 3 percentage points, respectively. The left panel suggests that these representation gaps are predicted to be larger than those for the last election (and thus for the current congressional session). Of course, this could be due to the poor performance of our model. Alternatively or additionally, this *may* be related to the impact of mid-decade redistricting, although we cannot conjecture anything based on our current predictions.

There are two caveats in this interpretation. First, our predictions focus on the four groups, and they cannot account for other groups in the population. Consequently, the above levels of over- and underrepresentation do not sum up to zero: the four groups do not exhaust the population categories. Second, we use the CVAP as a benchmark for the population share. Because Asians and Latinos tend to have lower rates of CVAP compared to their population shares, our predictions may be interpreted as a lower bound of racial descriptive representation.

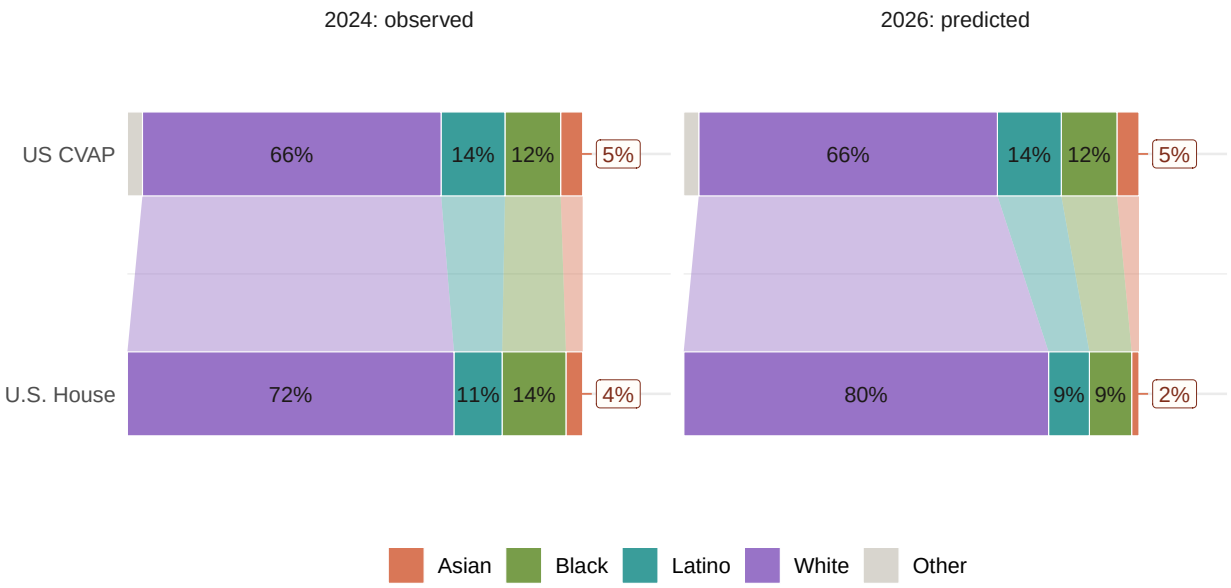


Figure 5: Difference between the Citizen Voting-Age Population and Our Predictions

5 How We Will Evaluate Our Predictions

Once official results become available, we will evaluate the accuracy of our predictions. To illustrate, we generate "hypothetical ground truth" data for racial descriptive representation. For each district, we assume that the plurality group (i.e., the group that has the largest share in CVAP) wins racial descriptive representation. Once we obtain actual ground truth data, we will simply replace our simulated data with it. We use multiple standard metrics for evaluating predictions and classifications (Kuhn, Johnson and Vaughan, 2013).

5.1 District Level

First, we assess our predictions at the district level with several standard metrics. For each racial group, we will leverage AUC-ROC, False Positive Rates, False Negative Rates, Sensitivity, Specificity, F1, Kappa, and the Brier score. Figure 6 shows our hypothetical results.

5.2 State Level

Next, we also assess our predictions at the state level. Here, we use visual confirmation as well as a few metrics, including the mean absolute error (MAE) and the root mean squared

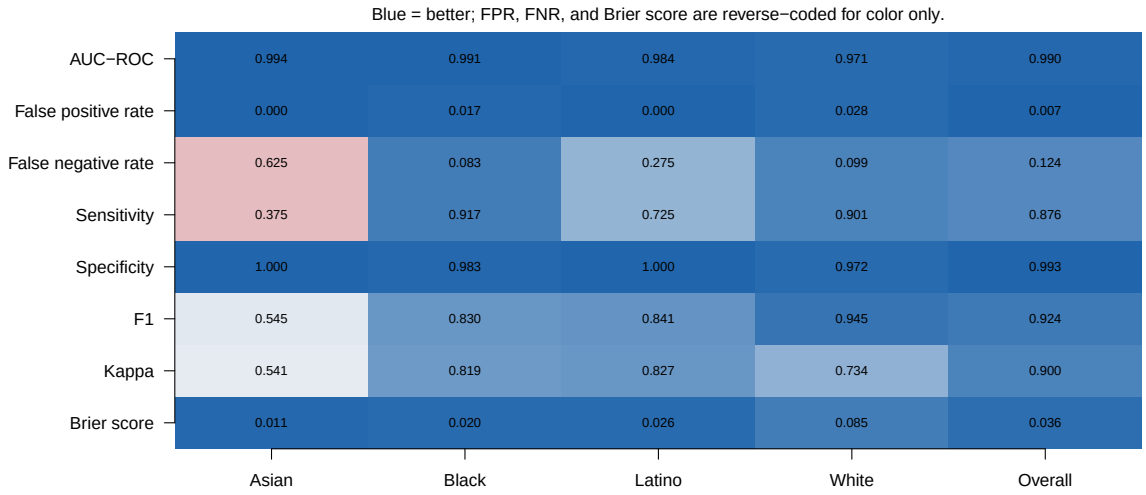


Figure 6: Predictive Performance at the District Level with Simulated Ground Truth Data

error (RMSE). Figure 7 presents our evaluations with hypothetical results.

5.3 National Level

Finally, we also assess our predictions for the entire legislature using a similar approach to the above. Figure 8 shows the predicted level of racial descriptive representation.

6 Additional Analysis

Beyond the above evaluation, we will perform several analyses. First, we will evaluate the predictive performance of our model for states with and without mid-decade redistricting, separately.

Second, we will train machine learning models to predict the same set of outcomes as our theoretical model. The goal is to compare the performance across the two approaches and assess to what extent the theoretical model, which does not learn (its parameter values) from the data, fares against the fully data-driven approach.

Third, we will compute the degree to which each district is "packed" with respect to the descriptive representation of a given racial group. We do so by computing the difference between the percentage of CVAP of a given group in a district and a predicted percentage of

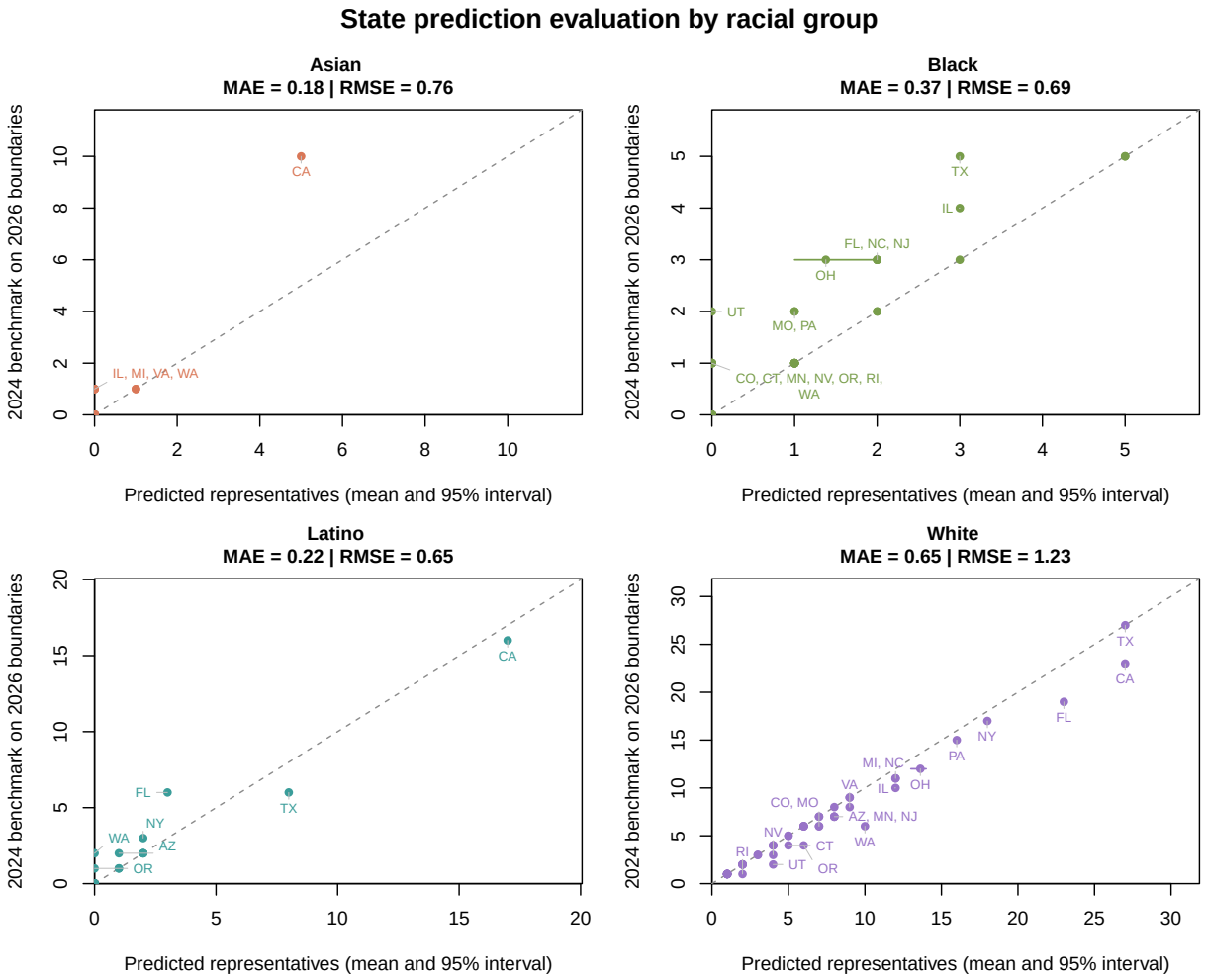


Figure 7: Predictive Performance at the State Level with Simulated Ground Truth Data

the same group that is predicted to produce a moderate to high chance of electoral victory, known as a "sweet spot" of districting (Lublin et al., 2019; Atsusaka, 2021).

Fourth, we will extend our predictions to the U.S. Senate elections. Predicting racial representation in the Senate is a unique challenge as district racial composition remains relatively stable (as opposed to cross-district variation for the U.S. House). Fifth, we will analyze state-level racial descriptive representation.

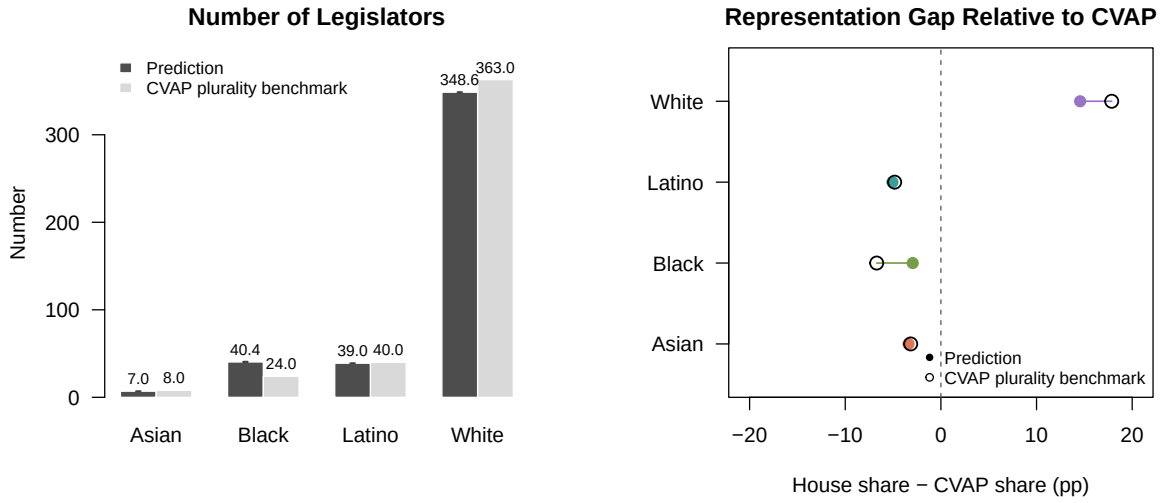


Figure 8: Predictive Performance at the National Level with Simulated Ground Truth Data

7 Timeline

- October 2026: Pre-analysis Plan
- November 3, 2026: Election Day
- November 2026: Data collection on observed candidate vote shares and race
- December 2026: Evaluating predictions
- January 2027: Additional analysis with machine learning
- February 2027: Complete the analysis

References

Abott, Carolyn and Asya Magazinnik. 2020. "At-Large Elections and Minority Representation in Local Government." *American Journal of Political Science* .

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.

Brace, Kimball, Bernard N Grofman, Lisa R Handley and Richard G Niemi. 1988. "Minority voting equality: The 65 percent rule in theory and practice." *Law & Policy* 10(1):43–62.

Bullock III, Charles S. 2010. *Redistricting: The most political activity in America*. Rowman & Littlefield Publishers.

Cameron, Charles, David Epstein and Sharyn O'halloran. 1996. "Do majority-minority districts maximize substantive black representation in Congress?" *American Political Science Review* 90(4):794–812.

Canon, David T. 2022. "Race and redistricting." *Annual Review of Political Science* 25:509–528.

Casellas, Jason P. 2011. *Latino representation in state houses and congress*. Cambridge University Press.

Davidson, Chandler and Bernard Grofman. 1994. *Quiet Revolution in the South: The Impact of the Voting Rights Act, 1965-1990*. Princeton University Press.

Epstein, David and Sharyn o'Halloran. 1999. "Measuring the electoral and policy impact of majority-minority voting districts." *American Journal of Political Science* pp. 367–395.

Fraga, Bernard L, Eric Gonzalez Juenke and Paru Shah. 2020. "One Run Leads to Another: Minority Incumbents and the Emergence of Lower Ticket Minority Candidates." *The Journal of Politics* 82(2):771–775.

Georgia v. Ashcroft. 2003. "Georgia v. Ashcroft." 539 U.S. 461. U.S. Supreme Court.

Grofman, Bernard, Lisa Handley and David Lublin. 2001. "Drawing effective miority districts: A conceptual framework and some empirical evidence." *NCL rev.* 79:1383.

Grofman, Bernard, Lisa Handley and Richard G Niemi. 1992. *Minority representation and the quest for voting equality*. Cambridge University Press.

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

Guinier, Lani. 1994. *The tyranny of the majority: Fundamental fairness in representative democracy*. Free Press New York.

Handley, Lisa and Bernard Grofman. 1994. "The impact of the Voting Rights Act on minority representation: Black officeholding in southern state legislatures and congressional delegations." *Quiet revolution in the South: The impact of the Voting Rights Act, 1965–1990* pp. 336–50.

Hankinson, Michael, Joseph Loffredo and Asya Magazinnik. 2026. "Assessing District Elections as a Remedy in State Voting Rights Acts." *Available at SSRN 7154498* .

Hicks, William D, Carl E Klarner, Seth C McKee and Daniel A Smith. 2018. "Revisiting majority-minority districts and black representation." *Political Research Quarterly* 71(2):408–423.

Juenke, Eric Gonzalez and Paru Shah. 2016. "Demand and supply: Racial and ethnic minority candidates in White districts." *Journal of Race, Ethnicity and Politics* 1(1):60–90.

Kenny, Christopher T, Brian Zhou, Tyler Simko and Kosuke Imai. 2026. "Weakening the Voting Rights Act reduces minority representation and electoral competition." *Christopher T. Kenny and Brian Zhou contributed equally* .

Kuhn, Max, K Johnson and D Vaughan. 2013. *Applied predictive modeling in rr*. New York: Springer.

League of United Latin American Citizens v. Perry. 2006. "League of United Latin American Citizens v. Perry." 548 U.S. 399. U.S. Supreme Court.

Lee, Diana Da In and Yamil Ricardo Velez. 2024. "Rising tides or political ripcurrents? gentrification and minority representation in 166 cities." *Urban Affairs Review* 60(3):956–982.

Lee, Diana Da In and Yamil Ricardo Velez. 2025. "Measuring descriptive representation at scale: Methods for predicting the race and ethnicity of public officials." *British Journal of Political Science* 55:e110.

Lublin, David. 1997. *The paradox of representation: Racial gerrymandering and minority interests in Congress*. Princeton University Press.

Lublin, David, Lisa Handley, Thomas L Brunell and Bernard Grofman. 2019. "Minority success in non-majority minority districts: Finding the "sweet spot"." *Journal of Race, Ethnicity and Politics* 5(2):275–298.

Lublin, David and Matthew Wright. 2024. "Diversity matters: The election of Asian Americans to US state and federal legislatures." *American Political Science Review* 118(1):380–400.

Schuit, Sophie and Jon C Rogowski. 2017. "Race, representation, and the voting rights act." *American Journal of Political Science* 61(3):513–526.

Shah, Paru. 2014. "It takes a Black candidate: A supply-side theory of minority representation." *Political Research Quarterly* 67(2):266–279.

Shaw v. Reno. 1993. "Shaw v. Reno." 509 U.S. 630. U.S. Supreme Court.

Shugart, Matthew S and Rein Taagepera. 2017. *Votes from seats: Logical models of electoral systems*. Cambridge University Press.

Taagepera, Rein. 2007. *Predicting Party Sizes: The Logic of Simple Electoral Systems*. Oxford University Press.

Taagepera, Rein. 2008. *Making Social Sciences More Scientific: The Need for Predictive Models*. Oxford University Press.

Thernstrom, Abigail M. 1987. *Whose votes count?: Affirmative action and minority voting rights*. Harvard University Press.