POLITICS

Forecasting the 2026 US Midterm Elections

Changing Maps and Outcomes: Gerrymander Potential and District Dissimilarity in Forecasting the 2026 Midterms

Jay A. DeSart, *Utah Valley University, USA*

Jay A. DeSart is Professor of Political Science at Utah Valley University. He can be reached at jdesart@uvu.edu

Congressional redistricting presents a particular challenge for forecasting U.S. House elections because changes to district boundaries can alter electoral outcomes in ways that are difficult to measure. My earlier state-level forecasting model addressed this problem using a measure of Gerrymander Potential based on changes in congressional apportionment and partisan control of the redistricting process. The wave of mid-decade redistricting preceding the 2026 election, however, demonstrates the limitations of a measure tied to the decennial redistricting cycle. This article develops District Dissimilarity, a continuous GIS-based measure that compares the geographic configuration of congressional district maps used in successive elections. Replacing Gerrymander Potential with District Dissimilarity produces only a modest improvement in overall model fit, but provides a more direct and theoretically defensible measure of congressional boundary change. Applied to the 2026 election, the revised model forecasts 241 Democratic seats compared with 238 using the original specification. A Monte Carlo simulation produces a mean forecast of 240.3 seats and a 95% prediction interval of 227 to 253 seats.

Prepublished Article

Forecasting congressional elections presents a unique challenge during the years immediately following redistricting. Every ten years, states redraw their congressional district boundaries in response to the results of the decennial census. Although the national political environment, state partisanship, and previous election results remain important predictors of House outcomes, redistricting introduces another source of uncertainty that is much more difficult to quantify. The challenge is not simply that district boundaries change. The challenge is determining whether those changes themselves alter the electoral landscape enough to influence the outcome of future elections.

Previously I attempted to account for this uncertainty by incorporating a measure of *Gerrymander Potential* into a state-level forecasting model of U.S. House elections (DeSart, 2023). The idea behind that variable was straightforward: Partisan

This is a “preproof” accepted article for *PS: Political Science & Politics*. This version may be subject to change during the production process.

DOI: 10.1017/S1049096526102911

gerrymandering is most likely to occur when one political party controls the congressional redistricting process at the same time a state must redraw its districts to accommodate a change in the number of congressional seats apportioned to the state. The resulting variable was intentionally simple. It identified those states where the opportunity for partisan mapmaking was greatest and incorporated that information into the forecasting model. Despite its simplicity, the variable proved to be statistically significant and improved the model's overall performance.

The conclusion of that article also acknowledged an important limitation. The Gerrymander Potential variable assumed that the greatest opportunity for partisan mapmaking occurred when reapportionment forced a state to redraw its congressional map. That assumption was reasonable, but it was not complete. Congressional district boundaries can change substantially even when the number of seats assigned to a state remains the same. Utah provided one example. Although the state did not gain an additional congressional seat after the 2020 Census, the Republican-controlled legislature substantially redrew the boundaries of the state's 4th Congressional District. The result was not an additional Republican seat in Congress, but rather a congressional map that significantly strengthened an existing Republican seat. At the conclusion of the earlier article, I suggested that a more refined measure of Gerrymander Potential would require an objective measure of the amount of change in congressional district boundaries from one election to the next.

When I wrote those words, I expected that issue to become important eventually. I did not expect it to become relevant almost immediately. The original model was developed for the first election following the 2020 Census, when congressional redistricting traditionally would have been expected to conclude. Instead, the years that followed saw an extraordinary continuation of congressional mapmaking. Courts ordered new congressional maps in several states. Legislatures revisited maps that had already been enacted. Congressional district boundaries continued to change well after the initial post-census redistricting cycle had ended. The limitation identified in the conclusion of the earlier article was no longer simply a theoretical concern. It had become a practical problem for anyone attempting to forecast House elections during the remainder of the decade.

The question then becomes how we should measure those changes. If the objective is to account for the potential electoral consequences of redistricting, identifying only the institutional conditions under which partisan mapmaking *could* occur is no longer sufficient. We also need a way to measure the extent to which congressional district boundaries were *actually* reconfigured. That is the purpose of this article.

Using geographic information system (GIS) techniques, I develop a continuous measure of congressional district boundary disruption by comparing the geographic similarity of

Prepublished Article

congressional district maps across successive elections. Rather than relying on a simple indicator of whether the political conditions for partisan mapmaking existed, the new measure captures the degree to which district boundaries actually changed. I then incorporate this measure into the state-level House forecasting model introduced in the earlier study, compare its performance with the original specification, and use the revised model to generate a forecast of the 2026 U.S. House elections.

The Original Model

This paper builds upon the state-level forecasting model I proposed for the 2022 U.S. House election (DeSart, 2023). Unlike most congressional forecasting models, which generate predictions either at the district level or for the aggregate national seat total, that model generated a forecast of the number of Democratic House seats won in each individual state. State-level forecasts were then aggregated to generate a national projection.

The model relied on a parsimonious set of predictors intended to capture both the national political environment and the political characteristics unique to each state. National-level variables included August presidential approval and a party-adjusted midterm dummy variable. State-level variables included the number of congressional seats apportioned to each state, a one-election lag of Democratic House seats won, Democratic presidential vote share in the previous election, state party control, and a new variable I named *Gerrymander Potential*. Using observations from every congressional election between 1970 and 2020, the model produced a strong statistical fit while generating a forecast that compared favorably with many other academic forecasting models for the 2022 election.

The primary methodological contribution of that article was the introduction of the Gerrymander Potential variable. The objective was not to identify whether partisan gerrymandering had occurred in any particular state, but rather to account for the circumstances under which the redistricting process was most likely to alter aggregate House election outcomes.

Strengths and Limitations of the Gerrymander Potential Variable

The Gerrymander Potential variable was intentionally simple. It identified states in which two conditions existed simultaneously: reapportionment had changed the state's number of congressional seats following the decennial census, and one political party exercised complete control over the congressional redistricting process. The variable was also party adjusted, taking on positive values when Democrats controlled

Prepublished Article

redistricting and negative values when Republicans controlled the process. States with divided government, independent redistricting commissions, or court-drawn maps were assigned a value of zero.

This approach had several important strengths. First, it reflected the institutional reality that opportunities for partisan line drawing are greatest when legislatures must accommodate a different number of congressional districts after reapportionment. Second, it distinguished between situations in which one party possessed unilateral control over the map-making process and those in which institutional constraints limited partisan influence. Finally, despite its simplicity, the variable achieved statistical significance in the forecasting model and generally moved state-level predictions closer to the observed 2022 election outcomes.

At the same time, the 2023 article acknowledged that the measure did not capture every circumstance in which partisan map drawing could influence election outcomes. Most importantly, it implicitly assumed that meaningful changes to district boundaries would occur primarily in states whose congressional delegations changed as a result of reapportionment. Although reapportionment undoubtedly creates additional opportunities for partisan line drawing, it is not a necessary condition. States can substantially redraw district boundaries even when the number of congressional seats remains unchanged. The discussion of Utah in the original article illustrated this limitation. Following the 2020 census, Utah retained four congressional districts, yet the Republican-controlled legislature substantially altered the boundaries of the Fourth Congressional District in ways that likely strengthened Republican electoral prospects without changing the state's overall apportionment.

The limitations of the original measure have become even more apparent in advance of the 2026 election. Unlike the traditional post-census redistricting cycle, several states are now engaged in mid-decade redistricting through legislative action, court decisions, or both. These developments create the possibility of meaningful changes to congressional district boundaries independent of reapportionment. Under these circumstances, a measure based solely on changes in congressional apportionment can no longer distinguish between states with substantial map revisions and those whose district boundaries remain largely unchanged.

Prepublished Article

Toward an Objective Measure of Boundary Change

The limitations of the Gerrymander Potential variable led me to suggest in the original article that a more refined approach would require “an objective measure of the amount of change in the district maps from one decade to the next” (DeSart, 2023, p. 603). The wave of mid-decade redistricting in advance of the 2026 election makes such a

measure even more important. Congressional maps can now change substantially between elections even when a state's number of apportioned seats remains unchanged.

The approach developed here addresses that problem by directly measuring how much a state's congressional district map changes from one election to the next. Rather than inferring the potential effects of redistricting from institutional conditions alone, I use GIS data to compare the actual district boundaries used in consecutive elections. This allows the measure to capture changes that occur both during the normal post-census redistricting process and at other points during the decade.

Prepublished Article

Measuring District Dissimilarity

The challenge is that congressional district maps cannot be compared simply by matching district numbers across elections. Districts are frequently renumbered, split, consolidated, or substantially redrawn during the redistricting process. Consequently, the problem is fundamentally geographic rather than administrative. The question is not whether District 3 in one election corresponds to District 3 in the next election, but rather how much of each district's geographic area remains intact after redistricting.

Because the objective is to compare the spatial correspondence between two sets of congressional districts, the revised measure relies on standard GIS polygon overlay techniques. Overlay analysis has long been one of the fundamental operations in geographic information systems because it provides a systematic means of identifying the spatial relationships among two or more geographic layers (Longley et al., 2015).

Overlay techniques have been applied to a wide range of research problems, including land-use suitability analysis and spatial decision making (Malczewski, 2006). In political science, GIS overlay has also been used to reconcile nonconforming census and electoral boundaries in order to combine demographic and election data for spatial analysis (Amos et al., 2017). The present study applies the same underlying GIS principles to a different problem by using polygon overlay to measure the geographic correspondence between congressional district maps from successive elections.

A more directly related application of GIS to congressional redistricting comes from Crespin (2005), who developed a continuous measure of district change following the 2000 redistricting cycle. His approach used census tract boundaries and population data to estimate the proportion of a district's population that was new to the district following redistricting. This makes population central to Crespin's measure because his primary concern was constituency change and its effect on the electoral advantage of incumbent members of Congress.

The measure developed here is intended to capture something somewhat different. My objective is to simply measure the extent to which the geographic configuration of a state's congressional map changes from one election to the next. For that reason, I use the geographic area shared by districts in successive maps rather than the population contained within those areas. This approach directly measures the amount of geographic change in a state's congressional map, but it does not distinguish between boundary changes affecting densely and sparsely populated areas. A population-weighted version of the measure could potentially provide additional information about the electoral consequences of those changes. Incorporating population therefore represents a logical next step in refining the measure and evaluating whether it further improves the forecasting model.

Calculating District Similarity

The first step is to identify the degree of overlap between every congressional district before redistricting and every congressional district after redistricting. GIS intersection operations partition the two maps into all overlapping polygons while preserving the identity of the districts that contribute to each overlap. This produces a complete set of pairwise overlap areas between districts from the two elections. Overlay-based comparisons of this type are widely used in GIS because they provide a direct measure of the spatial correspondence between alternative geographic partitions.

Each district from the earlier election is then compared with every district from the subsequent election. Rather than attempting to match districts by their numerical designation, the revised measure identifies the district sharing the greatest geographic overlap with each original district. Conceptually, this asks a simple question: *Where did most of this district go?* Selecting the maximum overlap allows districts to retain their geographic identity even when district numbers change or boundaries are substantially redrawn.

The overlap associated with that best match is divided by the total area of the original district to produce a similarity score ranging from 0 to 1. A value of 1 indicates that the district remained geographically unchanged, while values approaching 0 indicate increasingly extensive boundary revision. Normalizing overlap by the original district area follows the general GIS principle of comparing polygons by the proportion of shared area rather than by their absolute size, making similarity directly comparable across districts of very different geographic extent.

State-level similarity is calculated by summing the maximum overlap areas for all districts within a state and dividing by the total area of the state's congressional districts.

Prepublished Article

This produces a single similarity measure for each state's congressional map. Ultimately, a state's district similarity score is calculated with the following formula:

$$S_{t,t+1} = \frac{\sum_{i=1}^{N_t} \max_j (A_{ij})}{\sum_{i=1}^{N_t} A_i}$$

where:

- i identifies each district in the original map at time t;
- j identifies each district in the subsequent map at time $t+1$;
- A_{ij} is the geographic area shared by original district $i$and subsequent district j;
- A_i is the total geographic area of original district i;
- N_t is the number of districts in the original map;
- $\max_j(A_{ij})$is the largest overlap between original district $i$and any district in the subsequent map.

From Similarity to District Dissimilarity

Because the forecasting model is intended to measure the magnitude of map change rather than map stability, the state similarity score is converted into a dissimilarity measure by subtracting it from one. Under this transformation, a value of 0 indicates that district boundaries remained essentially unchanged from one election to the next, while values approaching 1 indicate increasingly extensive changes to the state's congressional map.

As expected, most observations have District Dissimilarity scores at or near zero because congressional district boundaries do not change from one election to the next in most states. Across the 56 years of data in this analysis, the variable ranges from 0 to 0.46, with a mean of .038. Scores slightly above zero are usually the result of minor technical adjustments to district boundaries or improvements in the GIS boundary files rather than meaningful changes to the districts themselves. The highest scores occur when states undergo substantial redistricting, whether through the normal decennial process or the less common mid-decade redistricting that motivated this study.

As with the original Gerrymander Potential variable, the measure is then adjusted for party control of the redistricting process. Changes occurring under unified Democratic

Prepublished Article

control receive positive values, while changes occurring under unified Republican control receive negative values. States in which redistricting was conducted under divided government, by an independent commission, or through judicial intervention receive a value of zero. The resulting variable, therefore, incorporates both the magnitude of congressional boundary change and the institutional context in which that change occurred.

This approach preserves the theoretical logic of the original forecasting model while replacing an indirect measure of redistricting opportunity with a direct measure of observed boundary change.

The Revised Forecasting Model

Having developed the district dissimilarity measure, the next step is to incorporate it into the original state-level forecasting model. The revised specification retains the overall structure of the original model while replacing the Gerrymander Potential variable with the new District Dissimilarity (Party Control Adjusted) measure. All of the remaining independent variables are identical to those used in the original model, allowing the effect of the revised measure to be evaluated without otherwise altering the model specification.

Table 1 presents the results of the revised model. The new District Dissimilarity (Party Control Adjusted) variable is positive and statistically significant ($p < .001$). This suggests that geographically larger changes in congressional district boundaries, when adjusted for partisan control of the redistricting process, are associated with greater changes in a party's representation in a state's congressional delegation, while controlling for the other variables in the model.

Prepublished Article

Table 1 – Revised State-level Model of U.S. House Election Outcomes, 1970–2024

Independent Variable	Unstandardized Regression Coefficient	sig
Total Seats Apportioned	.114	< .001
Democratic Seats (1 election lag)	.809	< .001
State Party Control	.134	.007
Democratic 2PPV previous presidential election	.012	.001
District Dissimilarity (Party Control Adjusted)	2.477	< .001
August Presidential Approval (Party Adjusted)	.015	< .001
Midterm Dummy (Party Adjusted)	-.430	< .001
Constant	-1.437	< .001

N = 1400
R^2 = .965
S.E. $_{y|x}$ = 1.106

The remaining independent variables continue to perform largely as expected. States with larger congressional delegations continue to elect more Democratic representatives, and Democratic representation in the previous election remains the strongest predictor of subsequent election outcomes. Democratic presidential vote share, state party control, presidential approval, and the midterm effect all retain the expected direction and remain statistically significant. Most importantly, replacing the original Gerrymander Potential variable does not meaningfully alter the relationships among the remaining predictors.

Although the district dissimilarity variable is statistically significant, an equally important question is whether replacing the original Gerrymander Potential variable affects the overall performance of the forecasting model. Table 2 compares the revised specification with the original model reported in DeSart 2023.

As Table 2 shows, the revised model performs nearly identically to the original specification. The R^2 increases slightly from .964 to .965, while the standard error of the estimate declines from 1.114 to 1.106. In addition, the revised model is estimated using data from two additional election cycles, increasing the sample from 1,300 to 1,400

Prepublished Article

state-election observations.

Table 2 – Comparison of the Original and Revised State-Level Forecasting Models

Model Statistic	Original Model (1970–2020)	Revised Model (1970–2024)
R^2	.964	.965
S.E. $_{y\|x}$	1.114	1.106
N	1,300	1,400

Note. Original model statistics from DeSart (2023); revised model statistics from the revised forecasting model.

Prepublished Article

These improvements in overall model fit are modest, but improving model fit is not the only consideration in developing a forecasting model. As Lewis-Beck and Tien (2008) argue, changes to forecasting models should be theoretically driven rather than simply attempts to improve statistical fit. The variables included in the model should therefore have a theoretical justification and measure the concepts they are intended to represent as directly as possible. The original Gerrymander Potential variable performed reasonably well statistically, but it was tied to the institutional conditions surrounding the normal post-census redistricting process and therefore could not account for substantial changes to congressional maps that occurred at other times during the decade. District Dissimilarity removes that limitation by measuring the maps actually used in consecutive elections.

The fact that this change produces only a modest improvement in overall model fit is not especially surprising. Most congressional maps do not change between elections, and the two measures often identify the same major post-census redistricting events. The advantage of District Dissimilarity is that when maps do change outside that normal cycle, the forecasting model can now account for those changes rather than treating the districts as though they had remained unchanged. The revised measure therefore preserves the performance of the original model while providing a more theoretically defensible way of accounting for changes to congressional district boundaries.

Forecasting the 2026 U.S. House Election

Generating a forecast of the 2026 U.S. House Election is a simple matter of plugging in the values for each of the variables for all 50 states. The revised model first generates a predicted Democratic seat total for each state. Because these predictions are produced by ordinary least squares regression, they are often fractional values (e.g., 3.74 or 7.18 seats). Since House seats are indivisible, each state forecast is rounded to the nearest whole seat. In addition, forecasts are constrained to each state's feasible range of outcomes. Predictions below zero Democratic seats are recorded as zero, while predictions exceeding the number of House seats apportioned to the state are reduced to that maximum value. The resulting state-level forecasts are then summed to produce the aggregate national forecast.

Missouri provides a particularly timely example of the advantage of measuring district change from one election to the next. The state adopted a new congressional map in 2025, which initially resulted in a nonzero District Dissimilarity value for the 2026 forecast. In September 2026, however, the Missouri Supreme Court ruled that the new map could not take effect while it remained subject to a referendum, leaving the congressional districts used in 2024 in place for the 2026 election. District Dissimilarity for Missouri is therefore zero in the final forecast. This illustrates an advantage of an election-specific measure of redistricting: it reflects the congressional map actually used in each election rather than assuming that district boundaries remain fixed between decennial redistricting cycles.

Using this procedure, the model incorporating the partisan control-adjusted District Dissimilarity variable predicts that the Democrats will win 241 seats in the U.S. House of Representatives in the 2026 election. For comparison, I also applied the original forecasting model to the same 2026 data. The original specification forecasts 238 Democratic seats. Thus, although the revised District Dissimilarity measure represents a substantially different operationalization of congressional boundary change, it produces only a modest change in the aggregate national forecast.

Because any point forecast is subject to uncertainty, I also estimated the distribution of plausible election outcomes using a Monte Carlo simulation. For each of 100,000 iterations, each state's forecast was allowed to vary randomly around its point estimate by drawing from a normal distribution with a standard deviation equal to the model's standard error of the estimate. The resulting simulated forecasts were then rounded, constrained to each state's feasible range of outcomes, and aggregated using the same procedure described above.

Table 3 summarizes the simulation results. The average simulated outcome is 240.3 Democratic seats, while the median outcome is 240 seats and the modal outcome is

Prepublished Article

241 seats. The close correspondence among these values indicates that the distribution of simulated outcomes is centered very near the point forecast. The simulations further produce a 95% prediction interval ranging from 227 to 253 Democratic seats. Of the 100,000 simulated elections, only 36 (0.036%) produced fewer than the 218 seats required for a Democratic majority in the House. This all suggests that while a Democratic majority is clearly the most likely outcome, the precise size of that majority remains uncertain.

Table 3 – Forecast of Democratic Seats in the U.S. House of Representatives, 2026

Forecast Statistic	Democratic Seats
Point Forecast	241
Monte Carlo Mean	240.3
Median	240
Mode	241
95% Prediction Interval	227–253

Note. The point forecast is the sum of the rounded and truncated state-level predictions from the revised model. Monte Carlo statistics are based on 100,000 simulated elections in which each state's forecast was allowed to vary normally around its point estimate using the model's standard error of estimate before rounding, truncation, and aggregation. The prediction interval reports the 2.5th and 97.5th percentiles of the simulated aggregate outcomes.

Prepublished Article

Although the revised specification produces a forecast that differs little from the original model, it does so using a more direct and objective measure of congressional boundary change. The revised model therefore preserves the forecasting performance of the original specification while improving the conceptual foundation of one of its principal explanatory variables.

Discussion

The Utah case that motivated this study exposed two different issues with the original Gerrymander Potential variable. The first was that it assumed congressional district boundaries changed only after each decennial census. As a result, it had no way of accounting for court-ordered redistricting or other mid-decade changes to congressional maps that ended up being a central feature of the 2026 election cycle. Once that limitation became apparent, it was difficult to justify continuing to treat all elections within a redistricting cycle as though they were being conducted using the same district boundaries.

The revised District Dissimilarity measure addresses that problem by treating each election as a separate event. Rather than assuming district boundaries remain constant throughout the decade, it directly compares the districts used in one election with those used in the next. If the map changes, whether because of the normal decennial redistricting process, a court order, or a special legislative session, the measure changes with it. In that sense, the revised measure is not tied to the redistricting cycle. It is tied to the districts voters actually use on Election Day.

One of the more interesting findings is that this conceptual improvement has relatively little effect on the overall forecast. The revised model predicts 241 Democratic seats in 2026, compared with 238 under the original specification, while producing virtually identical measures of overall model fit. That suggests the original forecasting model was reasonably robust despite the limitations of the Gerrymander Potential variable. The revised measure therefore represents less of a correction to the model than an improvement in how one of its key concepts is measured.

The Utah example also suggests that another challenge remains. Although the revised measure captures changes to district boundaries regardless of when they occur, it still does not directly measure the partisan consequences of those changes. A court may redraw congressional districts using a process intended to be politically neutral, yet the resulting map may still alter the electoral prospects of one or both parties. Utah may prove to be such a case. The court-ordered map adopted for the 2026 election creates a district that appears likely to be considerably more favorable to Democrats than any district under the overturned legislative map, although whether that change ultimately produces an additional Democratic seat will not be known until the election.

That limitation points toward the next step in improving state-level House forecasting models. This study moves beyond a measure tied to the decennial redistricting cycle by treating each election as a unique event. The next challenge is developing a way to measure not simply that district boundaries changed, but how those changes are likely to affect future election outcomes.

Prepublished Article

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the Harvard Dataverse at https://doi.org/10.7910/DVN/AQJTYI.

CONFLICTS OF INTEREST

The author declares no ethical issues or conflicts of interest in this research.

Prepublished Article

References

Amos, Brian, Michael P. McDonald, and Russell Watkins. 2017. “When Boundaries Collide: Constructing a National Database of Demographic and Voting Statistics.” *Public Opinion Quarterly* 81 (Special Issue): 385–400. doi: 10.1093/poq/nfx001.

Crespin, Michael H. 2005. “Using Geographic Information Systems to Measure District Change, 2000–2002.” *Political Analysis* 13 (3): 253–260. doi: 10.1093/pan/mpi016.

DeSart, Jay A. 2023. “A State-Level U.S. House Election Forecast Model for 2022: Modeling the Potential Effects of Gerrymandering.” *Polity* 55 (3): 511–535. doi: 10.1086/725244.

DeSart, Jay A. 2026. “Replication Data for: Changing Maps and Outcomes: Gerrymander Potential and District Dissimilarity in Forecasting the 2026 Midterms.” *PS: Political Science & Politics*. DOI: 10.7910/DVN/AQJTYI.

Lewis-Beck, Michael S., and Charles Tien. 2008. “Forecasting Presidential Elections: When to Change the Model.” *International Journal of Forecasting* 24 (2): 227–236. doi: 10.1016/j.ijforecast.2008.02.008.

Longley, Paul A., Michael F. Goodchild, David J. Maguire, and David W. Rhind. 2015. *Geographic Information Science and Systems*. 4th ed. Chichester, UK: John Wiley & Sons.

Malczewski, Jacek. 2006. “GIS-Based Multicriteria Decision Analysis: A Survey of the Literature.” *International Journal of Geographical Information Science* 20 (7): 703–726. doi: 10.1080/13658810600661508.