

A Compact Simulation Forecast of the 2026 General Election in Illinois' 9th Congressional District

Ryan McComb

rmccomb [at] votehub [dot] com

July 7, 2026

Abstract

This is a forecast of the 2026 general election in Illinois' 9th Congressional District, Biss (D) versus Ellison (R), built for fun in a single night with as little code as I could manage. The race is not competitive, but it is one that hits close to home, and the first forecast I built of this seat ran on Stata, slow and clunky; I wanted to see how few lines of Python could do the job properly. The answer came to 89 for vote share and 66 for turnout. Across 10,000 simulations Biss takes a median 73.9% of the two-party vote (80% interval: 71.3 to 76.2) on roughly 287,000 ballots.

1 Data

Everything runs on one precinct file, drawn on 2026 boundaries: 436 precincts spanning Chicago, suburban Cook County, and Lake and McHenry Counties, each carrying its 2024 presidential result and its census profile (race and ethnicity of the voting-age population, age, housing, density). The national environment starts from generic-ballot polling, which sits around D+7 as of writing. Rather than freeze that number, each simulation picks one of three Novembers: the environment keeps drifting toward the Democrats (D+9.5, 40% of the time), holds where it is (D+7, 35%), or tightens (D+4, 25%). The drift path gets the most weight because that is what midterms tend to do; Bafumi, Erikson and Wlezien [1] found that vote preferences move toward the out-party over the campaign year, by roughly five points of margin from February to November, about half of which remains from July.

2 Vote-share model

Every precinct starts where it finished in 2024, and then gets pushed three ways: a national swing everyone feels, a swing shared with precincts that resemble it, and a little noise of its own. The equation looks fancier than it is. With p_i as precinct i 's two-party Democratic share in 2024, each simulation draws

$$p_i^{2026} = \sigma \left(\text{logit}(p_i) + \frac{\delta_{\text{env}} + \delta_{c(i)} + \varepsilon_i}{25} \right), \quad (1)$$

where σ turns log-odds back into a share. The national swing δ_{env} is the drawn November environment plus some uncertainty; under the middle path it averages +4.9 points. The cluster swing $\delta_{c(i)}$ is shared within each of five precinct types (dense urban renter, heavily Hispanic, heavily Asian, inner suburb, outer suburb), which is the model's

way of admitting that when one heavily Asian precinct moves right, its cousins are probably moving too. The last term is precinct-level noise.

The one clever trick is that the swing is applied on the log-odds scale, which automatically makes safe precincts budge less. A flat swing would move an 88% precinct exactly as far as a 52% one, which is not how elections work; there is hardly anyone left to persuade in the first. On this scale the same national wave is worth about 5 points in an even precinct and roughly 2 in the district's most Democratic ones.

3 Turnout model

Turnout is two questions: how many ballots, and where. For how many, I regressed each precinct's 2024 turnout rate on its racial makeup, the old Goodman trick [3], which gives an implied turnout rate for each group. Each group then gets discounted by how much of its presidential turnout it typically keeps in a midterm, per the Census [2]: white voters keep about 81%, Black voters 72%, Hispanic voters 71%, Asian voters 67%. Those are national figures, and this district shows up to vote, so I flattened the differences most of the way toward uniform. Applied to each precinct's actual 2024 ballots, the district comes out at roughly 282,000 votes, 78% of its 2024 total. For where, an existing precinct-level turnout estimate decides how those ballots spread across the district; the electorate it produces leans slightly more Democratic than 2024's, which is what you would expect from a heavily college-educated seat. Turnout gets its own noise inside each simulation, and the district share is the turnout-weighted average of the precinct shares.

Worth saying plainly: that regression only knows about places, not people. The Hispanic and Asian rates it produces are artificially low in part because the denominator

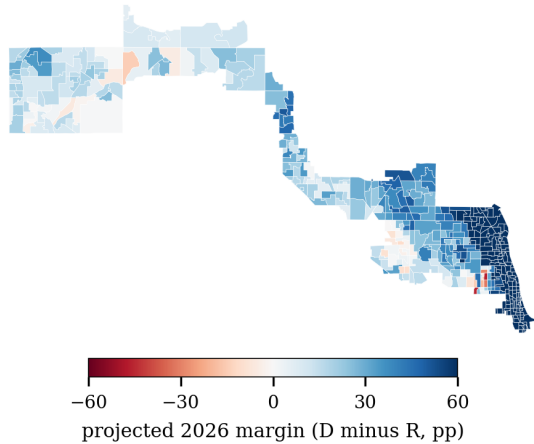


Figure 1: Projected 2026 two-party margin by precinct under the mean environment path: the lakefront wards project past D+60 while the McHenry County edge of the district stays within reach of even.

includes noncitizens. These numbers are turnout weights inside the model, not claims about how any actual group behaves.

4 Results

| Quantity | Estimate |
|----------------------------|----------------|
| Simulations | 10,000 |
| Biss win probability | >99.9% |
| Median two-party share (D) | 73.9% |
| 80% interval | 71.3% to 76.2% |
| Median margin | D+47.7 |
| Expected ballots cast | 286,765 |

Table 1: District topline across 10,000 simulations, environment paths mixed as described above.

Figure 1 maps the projection precinct by precinct. For comparison, the one published forecast of this race projects 189,892 Democratic and 72,807 Republican ballots, a 72.3% two-party share on 262,699 votes. This model lands about a point and a half above that on share and 9% above it on ballots. We disagree, albeit mildly, about how many people stay home in a midterm; we agree almost exactly about the shape of the district.

5 Limitations

The sizes of the three noise terms are educated guesses rather than measurements, because prior results do not exist on 2026 precinct lines; there is nothing to measure his-

torical precinct swings against, and building that crosswalk is the clearest path to a better-calibrated model. The five precinct types are threshold rules I wrote by hand. The retention ratios are national numbers applied to a district that likely drops off less than the nation. Candidate effects, third parties, and undervote are not modelled at all. King’s ecological inference method [4] is the proper version of the turnout regression, but it is a lot of machinery for 435 precincts, and several of the groups are thin slivers of most precincts’ population, which leaves the method little to work with. Each fix would cost lines of code, and the whole point was to have few of them.

References

- [1] Bafumi, J., Erikson, R. S., and Wlezien, C. (2010). Balancing, generic polls and midterm congressional elections. *The Journal of Politics*, 72(3), 705–719.
- [2] U.S. Census Bureau (2024). Voting and registration in the election of November 2022. Current Population Survey Voting and Registration Supplement.
- [3] Goodman, L. A. (1953). Ecological regressions and behavior of individuals. *American Sociological Review*, 18(6), 663–664.
- [4] King, G. (1997). *A Solution to the Ecological Inference Problem*. Princeton University Press.