

How Could the Census Citizenship Question Affect Congressional Apportionment and Gerrymandering?

Blog Source Model and Data

Our model is designed to simulate what proportion of districts a Republican candidate would be expected to win in a state, based on the following assumptions.

1. Voting data are available at the county level.
 - a. The ratio of Republican to Democrat votes for a state is the same as for the 2016 presidential vote (we ignored third-party candidates).
 - b. The ratio of Republican to Democrat votes for counties that voted Republican is the same as for the 2016 presidential vote (we ignored third-party candidates).
 - c. The ratio of Republican to Democrat votes for counties that voted Democrat is the same as for the 2016 presidential vote (we ignored third-party candidates).
2. Citizenship data are available at the county level, except for small-population counties where we estimate county citizenship based on the PUMA (Public Use Microdata Area) in which the small-population county is located.
 - a. Citizens as a percentage of the total population for a state is as estimated in the 2013–17 five-year American Community Survey (ACS).
 - b. Citizens as a percentage of the total population for counties that voted Republican is as estimated in the 2013–17 five-year American Community Survey (ACS).*
 - c. Citizens as a percentage of the total population for counties that voted Democrat is as estimated in the 2013–17 five-year American Community Survey (ACS).*
3. In our models, the voter turnout as a percentage of eligible voters is assumed to be the same for counties that voted Republican as for counties that voted Democrat.
4. In our models, the ratio of eligible voters to citizens is assumed to be the same for counties that voted Republican as for counties that voted Democrat.
5. In our models, the ratio of citizens to total population is assumed to be as calculated in step 2 above.
6. In our models for population-based intrastate districting, the population per district is assumed to be the same across counties that voted Republican as across counties that voted Democrat.
7. In our models for citizen-based intrastate districting, the number of citizens per district is assumed to be the same across counties that voted Republican as across counties that voted Democrat.
8. In a state gerrymandered for Republicans, we assume Republicans defeat Democrats in districts won by Republicans by a margin of 60 to 40 percent.
9. In a state gerrymandered for Republicans, we assume Republicans lose to Democrats in districts won by Democrats by a margin of 30 to 70 percent.

From these assumptions, we derive R , the expected proportion of districts that will be won by Republicans:

$$R = \frac{(c - b * \frac{e}{f})}{(a * \frac{d}{f} - b * \frac{e}{f})}$$

where

- R = the expected share of districts won by Republican candidates
- a = Republican votes as a share of Republican + Democrat votes in districts won by Republicans
- b = Republican votes as a share of Republican + Democrat votes in districts won by Democrats
- c = Republican votes as a share of Republican + Democrat votes, statewide
- d = citizens as a percentage of the total population in districts won by Republicans
- e = citizens as a percentage of the total population in districts won by Democrats
- f = citizens as a percentage of the total population, statewide

This relatively simple model has the following potential weaknesses:

- Our reliance on presidential votes means our models do not account for specific congressional races, candidates, or incumbency. This is the main factor that can create substantial differences between our model and the actual congressional votes in 2016.
- Our reliance on county-level voting data means our models miss some potential for districting based on voting patterns at the level of voting precincts, rather than whole counties.
- Our clustering of citizenship data into Republican-voting counties and Democrat-voting counties misses potential differences in citizenship levels by voting patterns among counties that voted for the same party, but our analyses did not find any systemic effects of this simplification.

Source Data

| | Texas | North Carolina |
|---|-------|----------------|
| a = Republican votes as a share of (Republican + Democrat) votes in counties won by Republicans | 67.9% | 65.2% |
| b = Republican votes as a share of (Republican + Democrat) votes in counties won by Democrats | 38.4% | 37.8% |
| c = Republican votes as a share of (Republican + Democrat) votes statewide | 54.8% | 52.0% |
| d = citizens as a percentage of total population in counties won by Republicans | 93.3% | 96.6% |
| e = citizens as a percentage of total population in counties won by Democrats | 85.1% | 93.2% |
| f = citizens as a percentage of total population statewide | 89.2% | 95.1% |

Information on arguments related to the fairness of citizen-based intrastate redistricting comes from Erika Lunder and Margaret Mikyung Lee. 2010, *Constitutionality of Excluding Aliens from the Census for Apportionment and Redistricting Purposes*, Washington, DC: Congressional Research Service.

County voting data come from Simon Rogers, "US Election 2016: How to Download County-Level Results Data," *Simon Rogers* (blog), November 16, 2016, <https://simonrogers.net/2016/11/16/us-election-2016-how-to-download-county-level-results-data/>.

Citizenship data by counties and by PUMAs for counties too small for public release come from Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, IPUMS USA: Version 9.0 American Community Survey 2013–17 five-year file, Minneapolis, MN: IPUMS, 2019, <https://doi.org/10.18128/D010.V9>.