

Red State, Blue State, Rich State, Poor State: Revisited

Re-examining American politics' most famous paradox in an era of shifting demographics and growing polarization



Why Americans Vote the Way They Do

Andrew Gelman

Overview

- **Introduction**
- **Project impetus**
- **Methodology**
- **Datasets**
- **Preliminary findings**
- **Conclusions and next steps**

Introduction: History of red, blue states

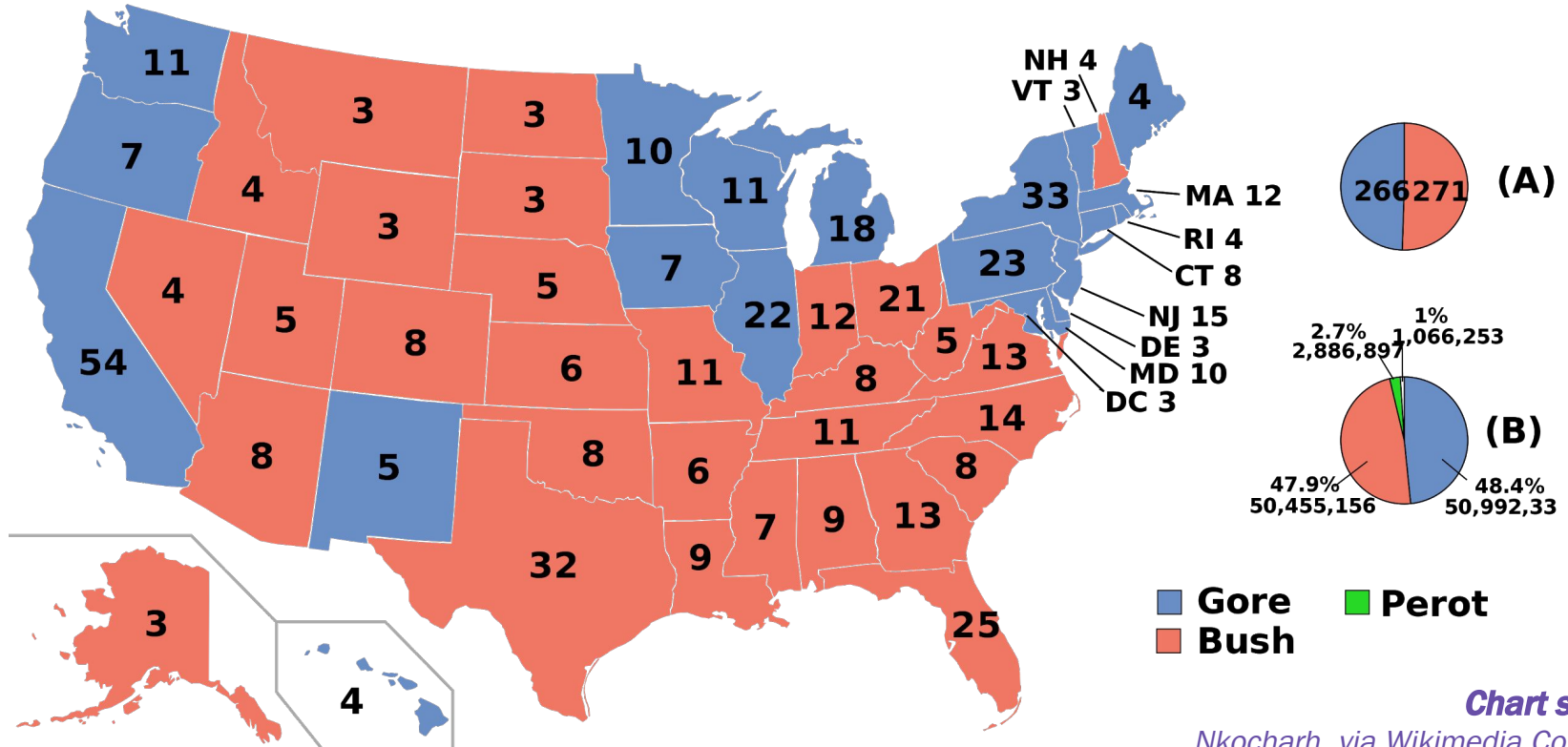
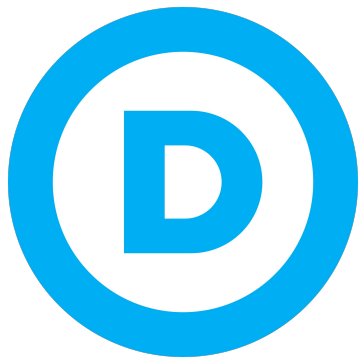


Chart source:
Nkocharh, via Wikimedia Commons

How did those colors get picked?

Until recent years, no official party colors in America.

- In Europe, red often used for left-leaning parties, blue for right-leaning.
- U.S. news outlets used own schemes, often echoing European one.
- **2000:** Just by chance, most outlets went red for GOP, blue for Dems.
 - Divisive, drawn-out, contested election cemented scheme in the public conscience.
- **2004:** Media outlets kept it. Each party used its color in its logos.



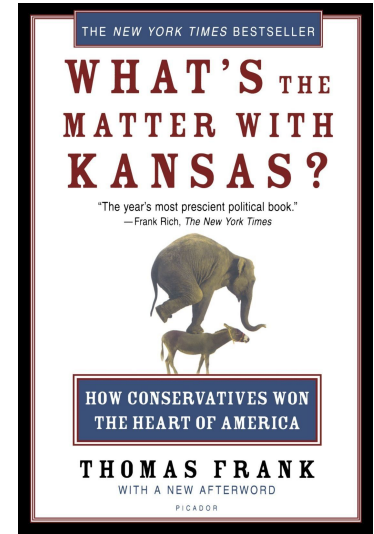
Enter the **red**-**blue** paradox

A clash of popular perceptions with data.

- **Perception:** Democrats as “rich, latte-sipping liberals,” Republicans as “poor, NASCAR-loving conservatives.”
- **Data:** Exit polls showed richer voters leaned GOP.

2005: Gelman, et al. address that clash:

- Affirmed exit polls’ findings that rich voters leaned **GOP**.
- But then, the researchers also noticed an oddity...



Rich states were trending **Democratic**, poor states **Republican**.

The paradox's causes and implications

What did it all mean?

- May explain “limousine Democrats” and “NASCAR Republicans.”
- Income's effect on vote choice more nuanced than previously thought.

So, can we reconcile/resolve the paradox?

- Gelman, et al.: Income-vote link weaker in rich states, stronger in poor states, even after accounting for demographics.
- Persisted even after adjusting for factors like black % of population.

Impetus: Where my study comes in

It's no longer 2004, and the political landscape has shifted.

- Red-blue map mostly steady in 2008, 2012, but not 2016.
 - Key Rust Belt “blue wall” states in flipped to Republicans.
 - Some states in Sun Belt showed signs of trending to Democrats.
- Emerging picture: Parties increasingly polarized along racial, urban-rural, educational, generational lines.
- **What about income? Conflicting signals.**

Goal: Build upon Gelman, et al.'s work by including 2008–2016 in order to assess red-blue paradox's present-day status.

Methods: Step-by-step overview

1. Replicate, expand state-level income-vote models

- How did a state's average annual income affect the share of votes it gave to the GOP candidate in every election since 1952?

2. Do same for individual-level income-vote models

- How did individuals' annual income affect their likelihood of voting for the GOP candidate in every election since 1952

3. Multilevel models on state, individual data

- Attempts to tie previous two together. Most challenging part.

Part 1: State-level model data

A few key datasets were needed:

1. States' average personal income: 1952–2016
 - Obtained from U.S. Bureau of Economic Analysis.
 - Values were adjusted for inflation (standard CPI), relative to 1996.
2. State-level presidential election returns: 1952–2016
 - Same as above, but 2008–2016 data obtained from U.S. Election Atlas.

Attempted to replicate these models, extend through 2016:

1. State's GOP candidate % vs. income via linear regression for each year.
2. Slope of each year's regression line vs. year via linear regression.

Part 1: State-level modeling

	Year	gel_koll0816_coef	gel5256_koll_coef
0	1952	0.069406	0.069406
1	1956	0.076889	0.076889
2	1960	0.006163	0.006572
3	1964	-0.146350	-0.146503
4	1968	-0.030032	-0.029460
5	1972	-0.070693	-0.064061
6	1976	0.044393	0.044516
7	1980	0.024031	0.023895
8	1984	-0.017204	-0.017041
9	1988	-0.030707	-0.029947
10	1992	-0.047100	-0.045655
11	1996	-0.070487	-0.067238
12	2000	-0.108180	-0.107851
13	2004	-0.104612	-0.106175
14	2008	-0.085777	-0.085777
15	2012	-0.084929	-0.084929
16	2016	-0.115505	-0.115505

Table: Linear regression model slopes.

- **‘Coef’**: beta coefficient. (Slope of regression line.)
- Self-calculated/-compiled income and electoral data yielded similar slopes as in Gelman, et al.
- 2008-2016: Values are totally new, and thus the same in both sets.
- 1952-1956: Same in both sets, as inflation adjusted data couldn't be calculated for 1952 and 1956.

These results underlie the following graphs...

Part 1: State-level modeling

Exact replica of Figure 1 from *Gelman, et al.* shown at right:

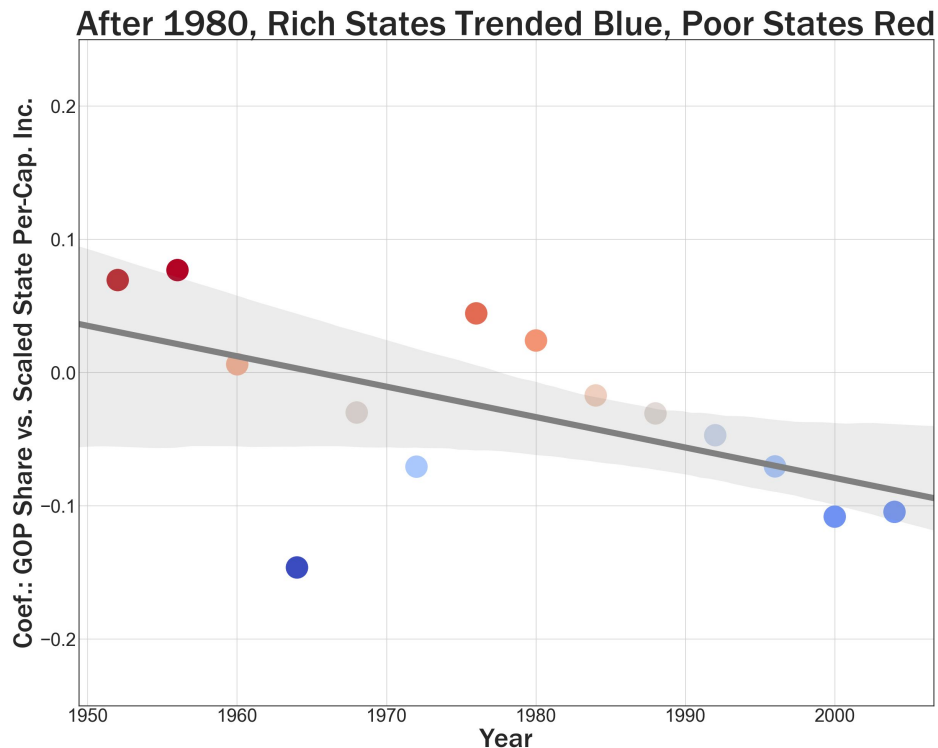
- Dot color → Rich states' party lean.
- Dot shade → Strength of that lean.

Pre-1980: No clear pattern.

- Wide confidence interval (95%).
- 1964: **LBJ's** 23-pt. rout over **Goldwater**

Post-1980: Clear trend.

- Rich states became more likely to support Democratic candidate.
- The reverse for poor states.

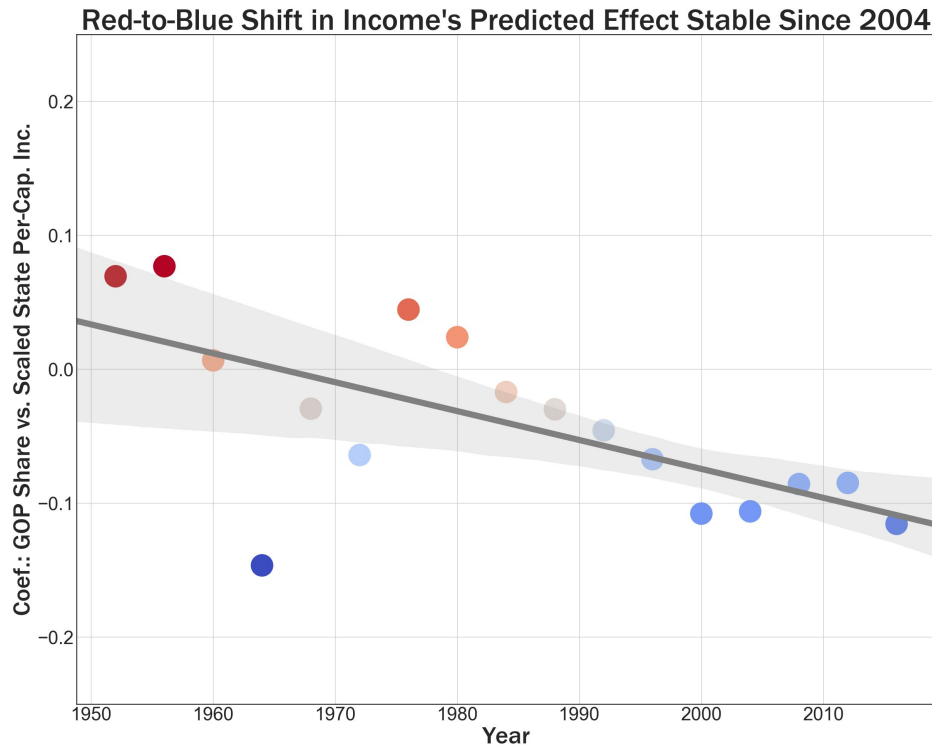


Part 1: State-level modeling

Adding 2008-2016:

- Mostly steady, with a slight wobbliness since 2004.
- Uptick in 2008, 2012: Possible artifact of the Great Recession, or just statistical noise?

Next steps? Consider role of other demographic, socioeconomic, geographical (urban-rural, region) factors in affecting this relationship.



Part 2: Individual-level model data

1. Individually paired income and vote data: 1952–2016
 - Gelman, et al.'s replication package had data for every year through 2004.
 - Data for later years obtained from same source: Bureau of Economic Analysis.
2. State-level presidential election returns: 1952-2016
 - Gelman, et al.'s replication package had data for every year through 2004.
 - Data for later years obtained from same source: Bureau of Economic Analysis.

Exact modeling approaches to be determined.

Conclusions and next steps

Preliminary conclusions:

- Attempts to replicate *Gelman, et al.*'s state- and individual-level income-vote models, graphs largely successful.
- Hypothesis to test: 2008–2016 shows paradox's effect waning.
- More data adjustments needed before hypothesis test possible.

Next steps:

- Continue to refine individual-level models, complete multilevel models.
- Consider breakdowns by race, ethnicity, education, age, region, urban-rural, other key demographic and socioeconomic factors.