Forecasting the 2020 US Elections

C. Lewis-Beck, M. Lewis-Beck

University of Iowa

November 5, 2020

Outline

- Introduction
- Forecasting Approaches
  - Political Economy Model
  - The Economist Model
  - Citizen Forecasting
- 2020 Preliminary Election Results
- Discussion

Model Considerations

- Accuracy: How close is the prediction to the actual outcome?
- Lead time: How early is the forecast made?
- Parsimony: Model should be realistic, yet also interpretable and concise.
- Transparency: Other researchers can produce the same forecast.

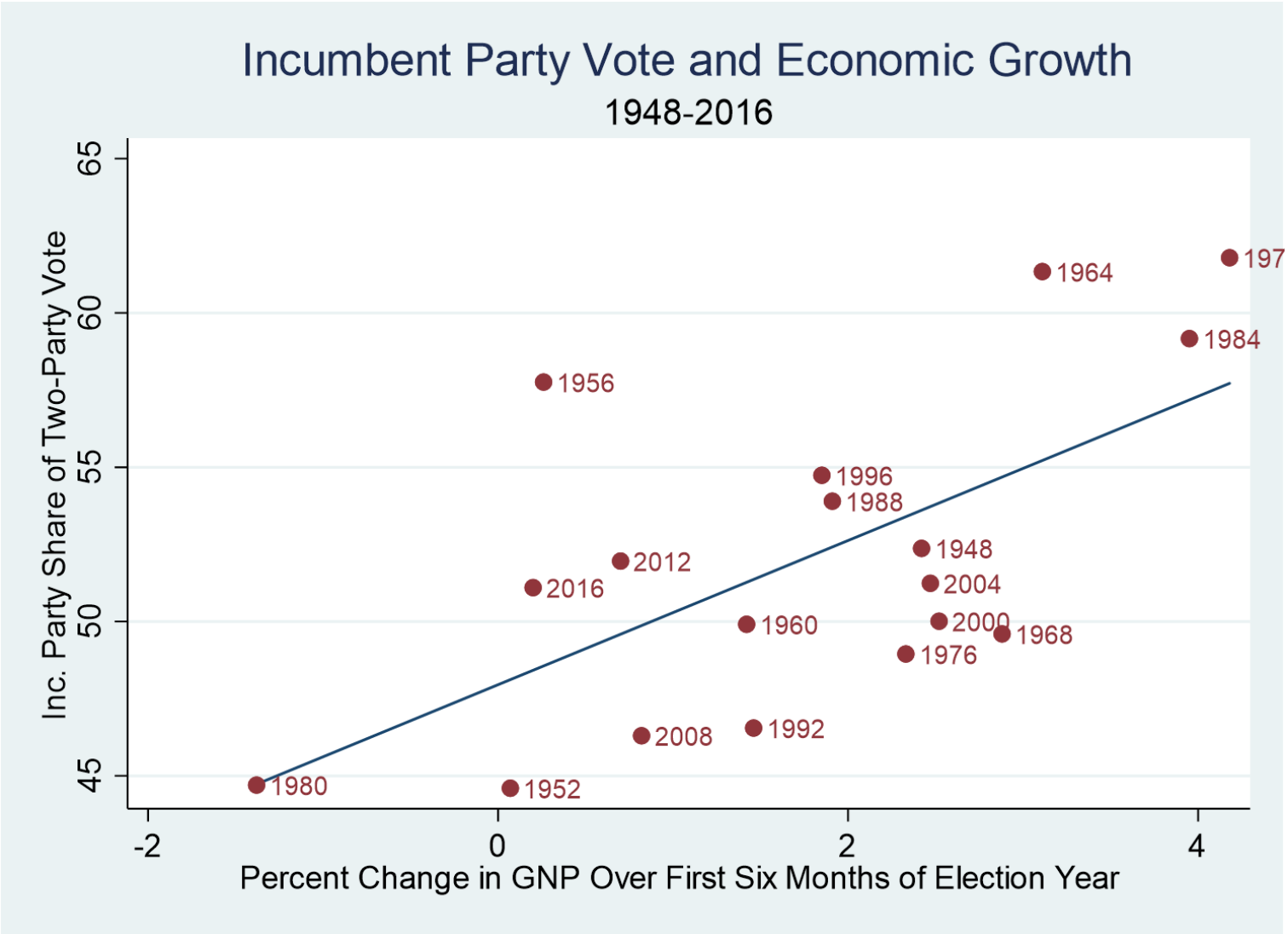
Political Economy Model

- One of the earliest political science forecasting methods.
- Theory driven based on fundamental political and economic issues.

$$\text{Incumbent Vote} = \text{Presidential Popularity} + \text{Economic Growth} + \epsilon$$

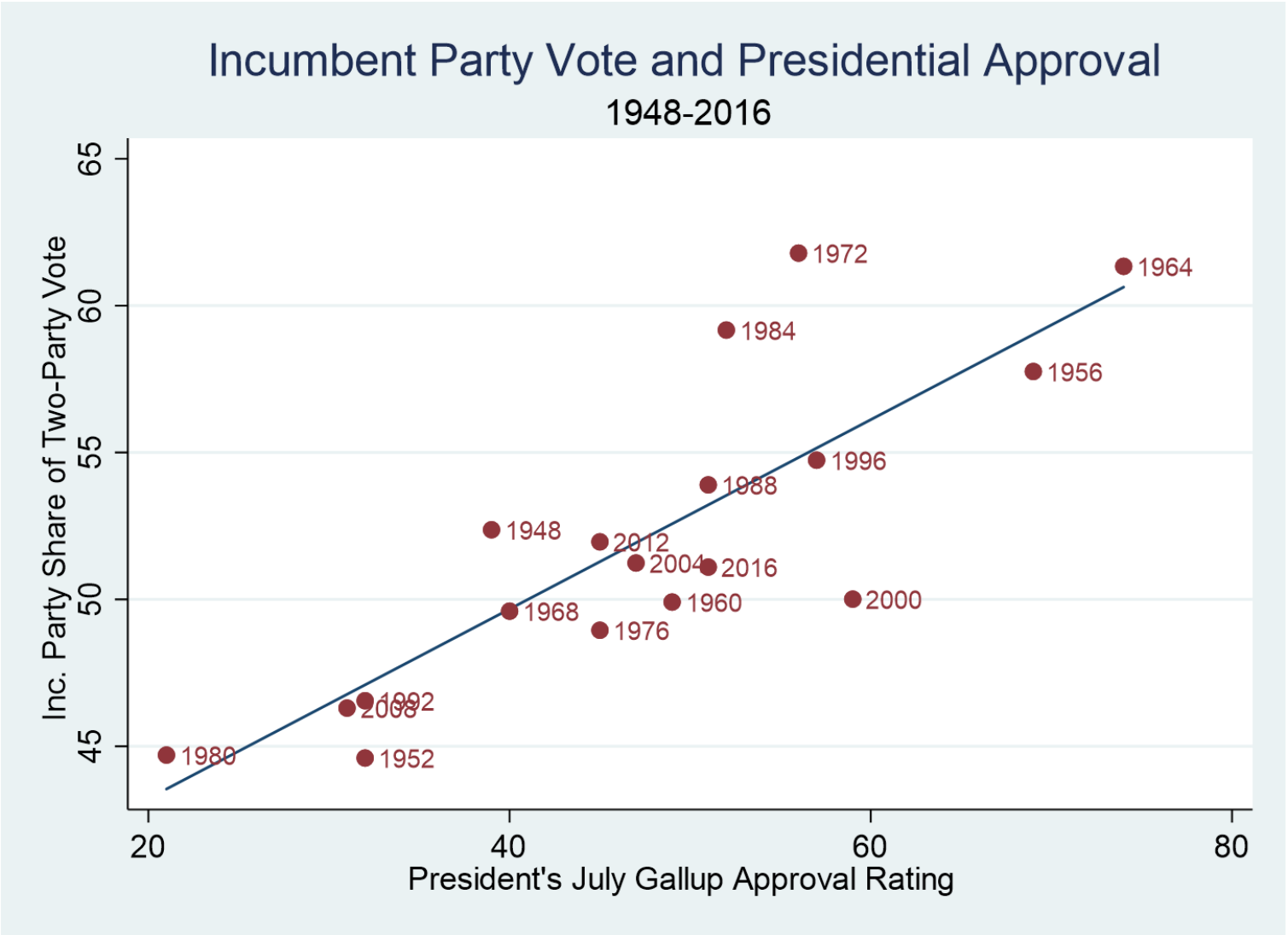
- Estimated well in advance of the election.

Political Economy Model



$r = 0.64$

Political Economy Model



$r = 0.82$

Political Economy Model

2020 Political Economy Model based on available data on August 27th.

$$\text{Popular Vote} = \underset{(15.37)}{37.50} + \underset{(4.73)}{0.26}\,\text{Pres Approval} + \underset{(2.25)}{1.18}\,\text{GNP} \quad R^2 = 0.76.$$

- Pres Approval based on Gallup data (July) $= 41$.
- GNP based on change over the first two quarters of the election year (Aug) GNP $= -4.14$.
- 2020 Forecast $= 43.3$.

Political Economy Model

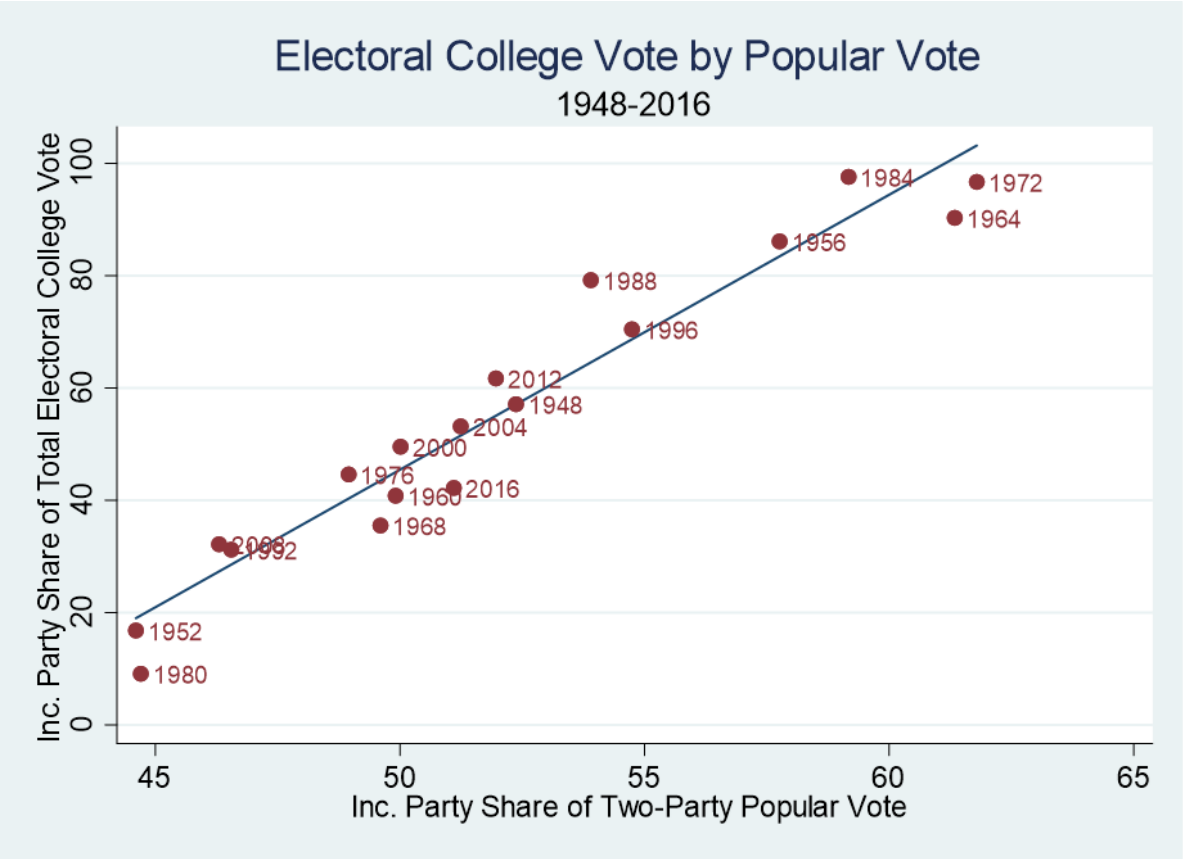
Model has been accurate for the two-party popular vote in the last ten elections.

Year	Two-party vote	Out-of-sample forecasts	Correct?
1948	52.4	50.4	Yes
1952	44.6	46.3	Yes
1956	57.8	54.2	Yes
1960	49.9	52.3	No
1964	61.3	60.4	Yes
1968	49.6	51.8	No
1972	61.8	55.8	Yes
1976	49.0	52.4	No
1980	44.7	39.1	Yes
1984	59.2	55.0	Yes
1988	53.9	53.2	Yes
1992	46.6	47.8	Yes
1996	54.7	54.7	Yes
2000	50.0	56.8	Yes
2004	51.2	52.9	Yes
2008	46.3	46.7	Yes
2012	52.0	50.0	Yes
2016	51.1	51	Yes

Political Economy Model

Predicting the Electoral College (EC) vote from the popular vote. Popular Vote of 43.3% gives EC forecast of 68 for Trump.

$$\underset{}{\text{EC Vote}} = \underset{(11.64)}{-199.42} + \underset{(14.94)}{4.90}\ \text{Popular Vote} \quad R^2 = 0.93.$$



Political Economy Model

Limitations

- Small sample size.
- Selection of lead time.
- Measurement of economic growth in 2020.

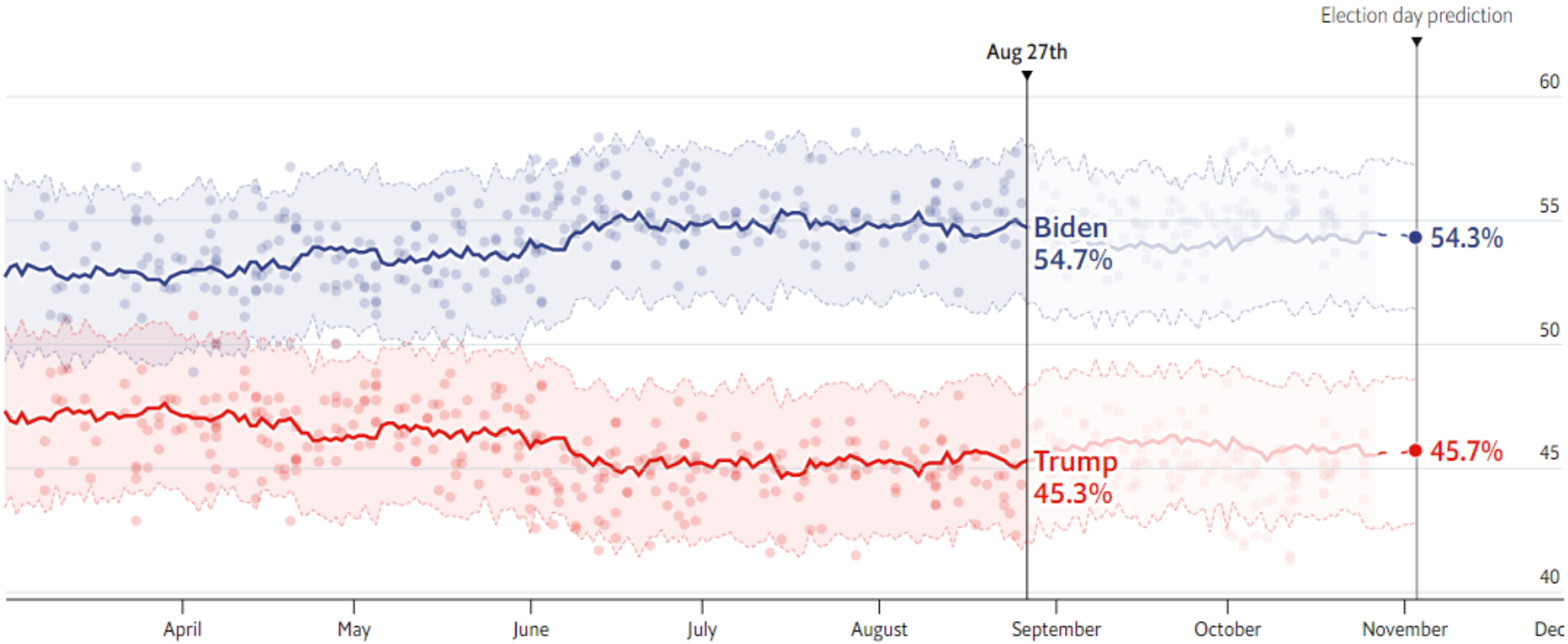
The Economist Model

The Economist Model (E-Model) combines a structural component with state and national polls.

- Estimated using a Bayesian framework similar to model developed by Linzer (2013).
- Structural component functions as the prior belief about final election results.
- Forecast updated daily based on new polling data.
- Incorporates a hierarchical structure to borrow information across states and time, which helps fill in gaps from missing polls and smooths temporal variability.

The Economist Model

Daily forecasts with uncertainty available on the Economist website.



The Economist Model

Features of the E-Model

- Combines historical regression-based models with real time polling data to improve accuracy.
- Tracks current voter preferences in addition to providing forecasts.
- Adjusts for polling biases (e.g., partisan nonresponse, quality of polling firm, etc.).
- Model is continually monitored and changed over time to modify forecasts and uncertainty bounds.

The Economist Model

Limitations

- Model specification is opaque. On August 5th, "We fixed a series of medium-sized bugs in the model that collectively caused too little uncertainty in our forecast..."
- A lack of a nontrivial lead time in forecasts.
- Emphasis on data mining/curve fitting rather than a structural/mechanism approach.

Citizen Forecasting

Query on voter expectations: "Who do you think will win the upcoming election?"

- Based on Condorcet's Jury Theorem on the "wisdom of the crowd."
- Successfully forecasted national elections in the United Kingdom (Murr, 2016; Lewis-Beck & Stegmaier, 2011).
- Yet to be tested *ex ante* in the United States at the state level.

Citizen Forecasting

Data collected from each US state in July using Amazon's Mechanical Turk.

- Question 1: "Who do you think will be elected President in November?"
- Question 2: "Which candidate for President do you think will win in this state?"
- Total sample size, $N = 2,483$. Median sample size within each state, $N = 60$.

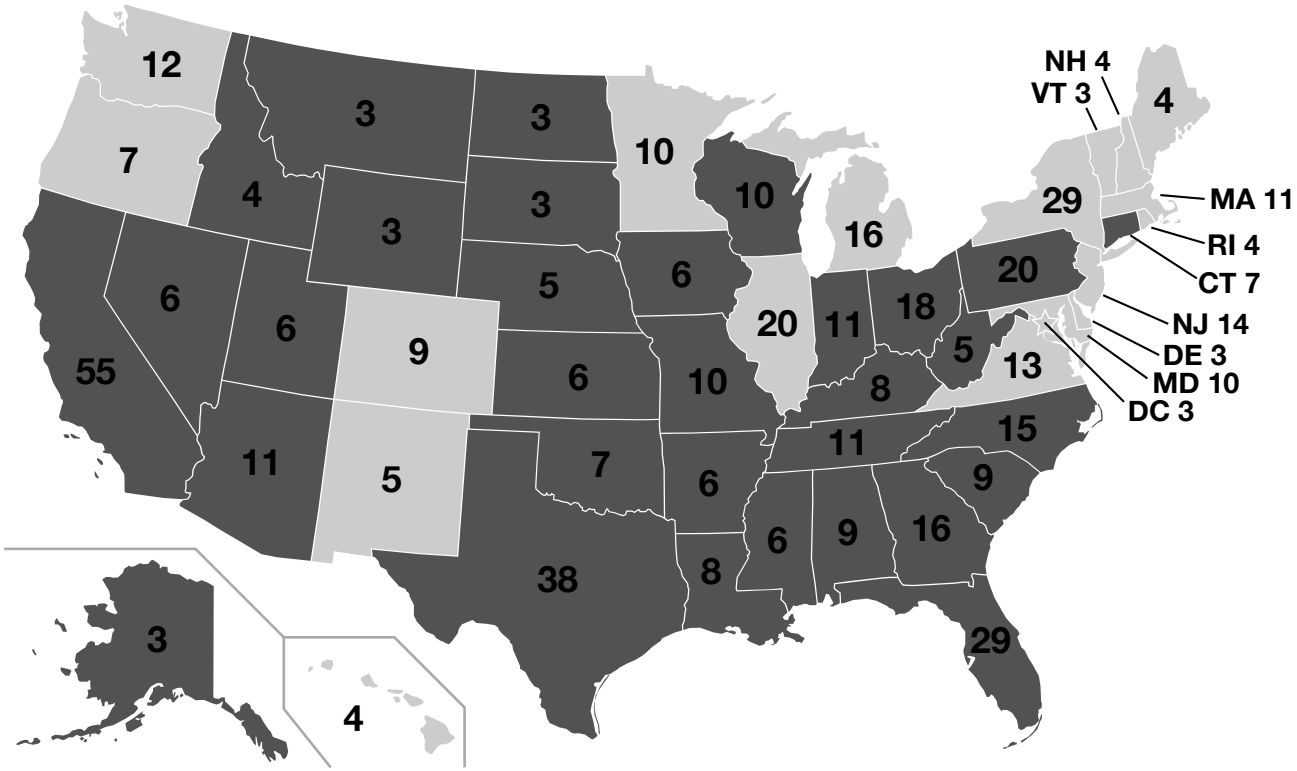
Citizen Forecasting

Results

- In the nation as a whole, 51.35 percent of individuals believe Joe Biden will win.
- Question 2, however, tells a different story
- Declare a state for Biden if over 50 percent of state respondents in that state call it for him

Citizen Forecasting

Biden 204 electoral votes, Trump 334 electoral votes



Trump = Dark Grey, Biden = Light Gray

Citizen Forecasting

- Is the quality of the MTurk sample off?
- Characteristics of the sample suggest they are engaged in politics (2.5 on 3.0 scale).
- Survey respondents have an approval rating of over 90 percent on Amazon Turk.
- MTurk national forecast (as measured by voter expectations) yields a forecast essentially equal to the national citizen forecast from *The Economist/YouGov Polls* (where 39% expected a Biden win and 40% a Trump win as of July 12-14 (remaining were undecided)).

Election Results (as of noon, November 5)

Total popular vote share:

$$\text{Biden} = 72,401,815 = 50.5\%$$
$$\text{Trump} = 68,606,127 = 47.8\%$$

Two-party popular vote share:

$$\text{Biden} = 51.35\%$$
$$\text{Trump} = 48.66\%$$

PE Model

Problems measuring the economy and presidential popularity

- Adjustment to the Economic Growth variable (Winsorizing it to equal the most extreme negative growth rate in the time series of the test data, i.e., -1.38).
- The 95% CI for Trump's Electoral Vote share then becomes (78,230).
- Trump's unusual popularity pattern.

E Model

"If your experiment needs a statistician, you need a better experiment." *Ernest Rutherford*.

- State and national polls are still biased.
- E-model relies more heavily on polling data as election approaches.
- E-model adjustments did not correctly fix the polling data.

Ensemble Model?

- The Citizen Forecasting model Electoral College estimate, at 334 for Trump, is too high.
- The Political Economy model has a similar problem, but in the other direction, a landslide for Biden.
- An Ensemble forecast, based on averaging forecasts from the two models, namely the PE forecast and the CF forecast, yields, $(68 + 334)/2 = 201$ electoral votes for Trump (closer to his current share).

References

- Colin Lewis-Beck and Michael Lewis-Beck, *U.S. Presidential Election Forecasting: The Economist Model*, Foresight: The International Journal of Applied Forecasting (2020), no. 59, 38–44.
- Michael S Lewis-Beck and Mary Stegmaier, *Citizen forecasting: Can UK voters see the future?*, Electoral Studies **30** (2011), no. 2, 264–268.
- Michael S Lewis-Beck and Charles Tien, *The political economy model: A blue wave forecast for 2020*, PS: Political Science & Politics (2020), 1–4.
- Drew A Linzer, *Dynamic Bayesian forecasting of presidential elections in the states*, Journal of the American Statistical Association **108** (2013), no. 501, 124–134.
- Andreas E Murr and Michael S Lewis-Beck, *Citizen forecasting 2020: A state-by-state experiment*, PS: Political Science & Politics (2020), 1–5.
- Andreas E Murr, *The wisdom of crowds: What do citizens forecast for the 2015 British General Election?*, Electoral Studies **41** (2016), 283–288.