The Elusive Likely Voter

Improving Electoral Predictions with More Informed Vote Propensity Models

Anthony Rentsch, Harvard University
Brian Schaffner, Tufts University
Justin Gross, University of Massachusetts Amherst

AAPOR 2019
Concurrent Session J

Motivation

- Nearly all horse race polls report estimates in terms of likely voters
- Few make the specifics of their likely voter model publically available
  - Makes it hard to systematically study their effects in the wild!
- **Goal**: develop a likely voter model framework that
  - improves horse race estimates
  - is easy and straightforward for others to implement

Data

Cooperative Congressional Election Study (CCES)

- Vote validation using Catalist's voter file
- Large N across states and years
- Measures standard demographic and attitudinal variables that are correlated with turnout and mis-reporting

Modeling approach

Follow framework in Keeter, Igielnik, and Wiesel, 2016.

1. Cutoff approaches
    a. Self-reported vote intent
    b. Re-formulated Perry-Gallup index
2. Probabilistic approaches
    a. **Inputs**: Perry-Gallup index variables
    b. **Inputs**: Perry-Gallup index variables + demographic variables

For (2), we test both logistic regression and random forests.

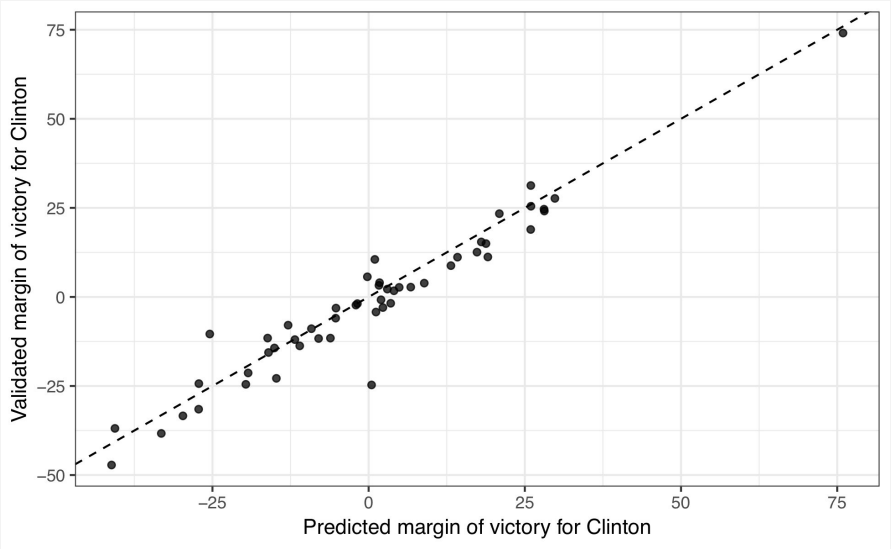
2016 national results

Approach	Implied Turnout	National Bias	Avg. Bias by State	Avg. Absolute Error by State
		Cutoff Approaches		
Already voted + will definitely vote	70.78%	3.59	2.46	4.44
Perry Gallup 6's	41.66%	-1.70	-2.76	6.19
Perry Gallup 6's + 5's	60.26%	2.05	1.18	3.98
		Probabilistic Approaches		
Perry Gallup	66.55%	3.29	1.75	4.17
Perry Gallup + Demographics	59.86%	-0.19	-0.36	4.02

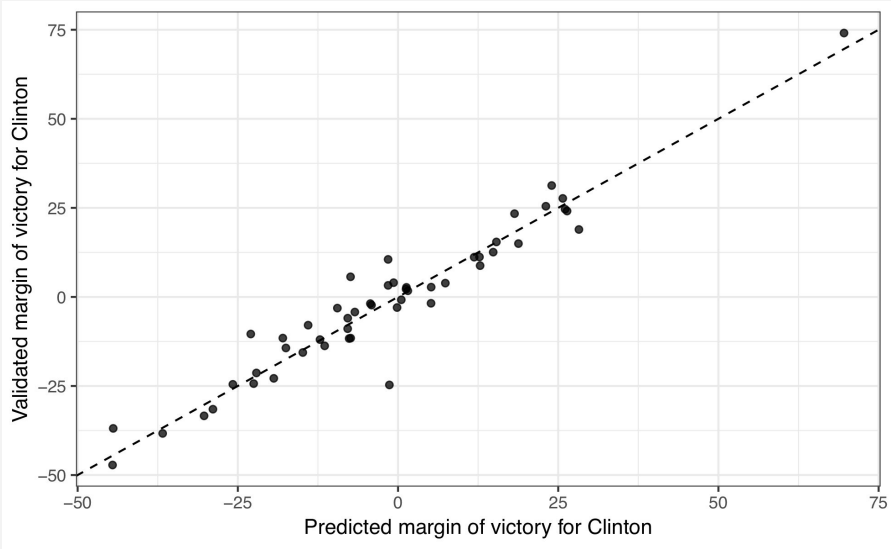
Validated turnout among 2016 CCES respondents: 55.11%

2016 state results

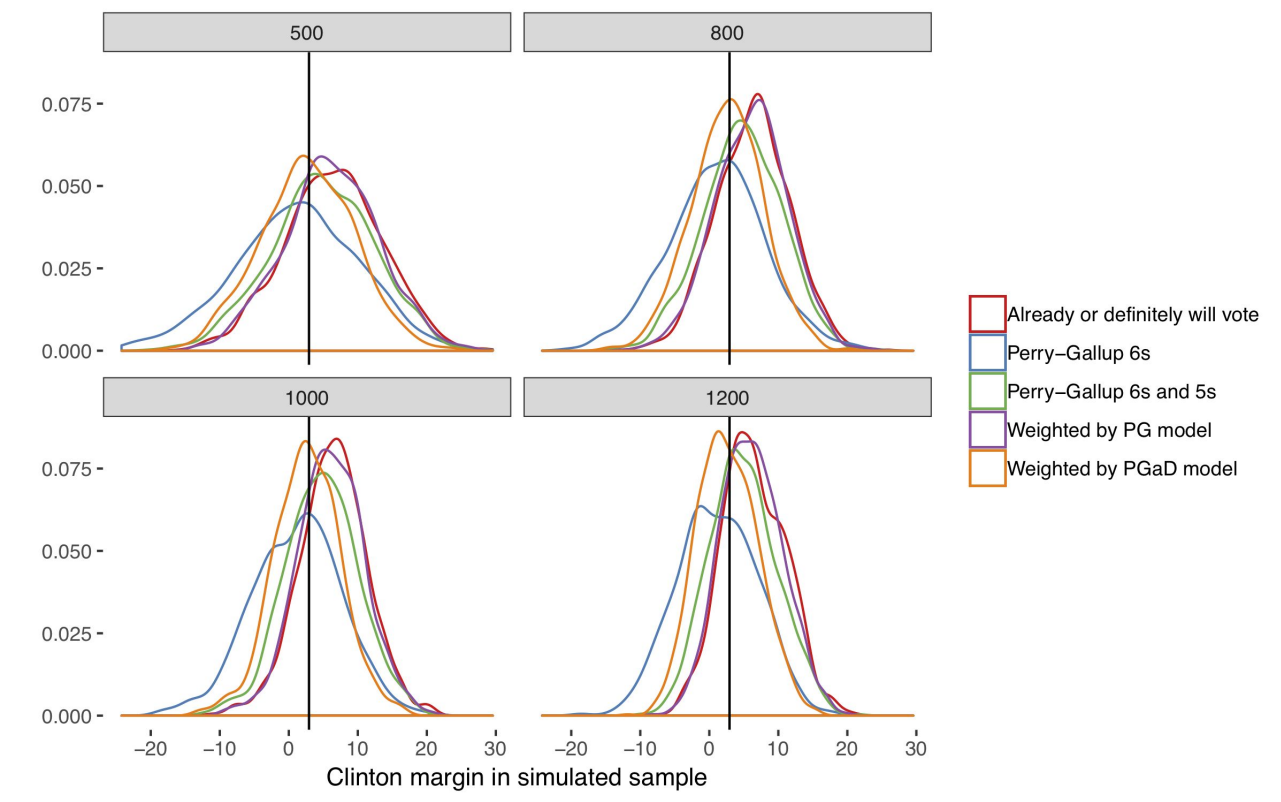
Perry-Gallup



Perry-Gallup and Demographics



2016 simulations



2014 national results

Approach	Implied Turnout	National Bias
	Cutoff Approaches	
Already voted + will definitely vote	73.01%	2.41
Perry Gallup 6's	44.26%	-5.06
Perry Gallup 6's + 5's	66.35%	1.20
	Probabilistic Approaches	
Perry Gallup	75.11%	2.42
Perry Gallup + Demographics	69.89%	0.50

Validated turnout among 2014 CCES respondents: 50.1%

Discussion and future directions

- Probabilistic approach with Perry-Gallup and demographic variables performs best
  - Reduces bias
  - Stable across years and smaller samples
- Make model/source code available
- Possibility of reporting estimates with respect to different turnout scenarios

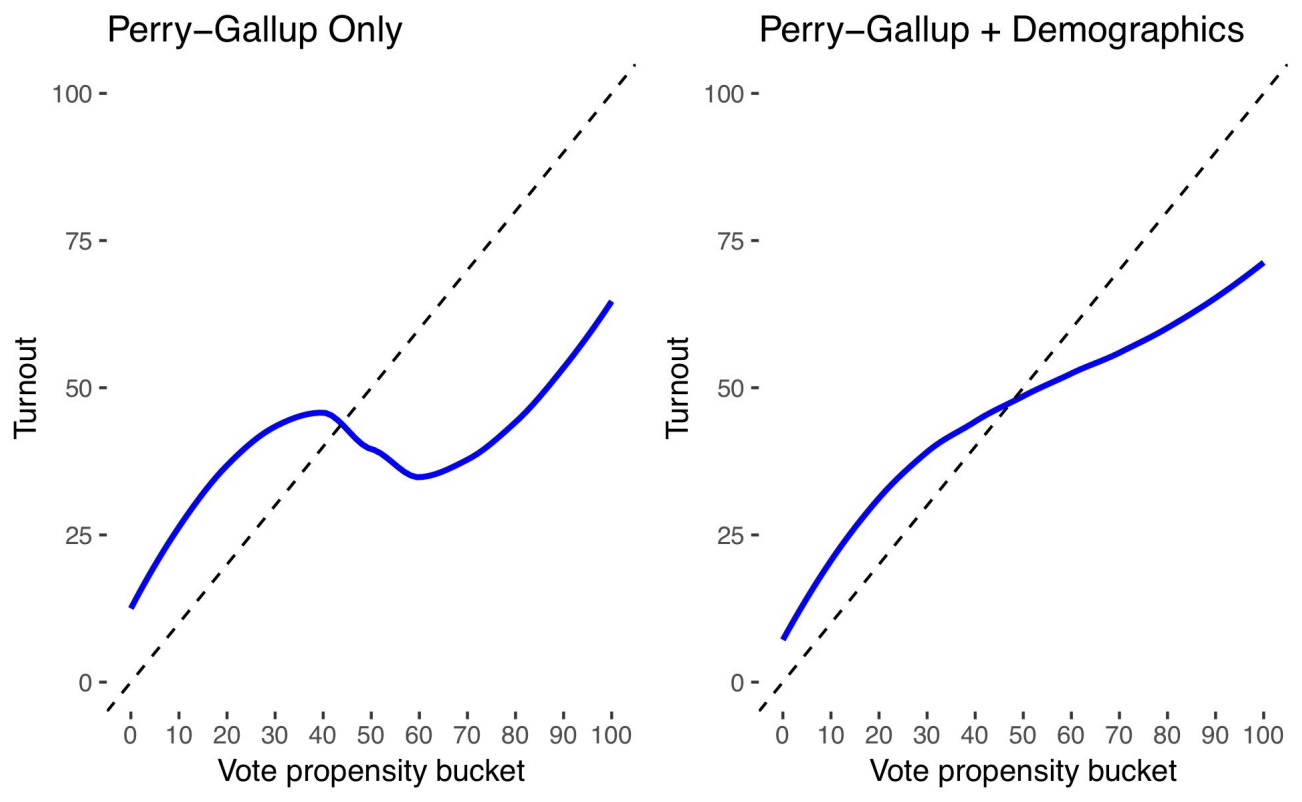
Thank you!

Anthony Rentsch

anthony.rentsch@g.harvard.edu

@Anthony_Rentsch

2016 model validation



2014 model validation