INSTITUTE FOR SOCIAL RESEARCH • SURVEY RESEARCH CENTER
SURVEY RESEARCH OPERATIONS
UNIVERSITY OF MICHIGAN

How Much Measurement Error in Recalled Vote Can Offset Nonresponse Error Reductions in Weighting? A Simulation Study

Raphael Nishimura

AAPOR 80th Annual Conference, St. Louis, May 15

Introduction

- Nonresponse error is one of the major threats in surveys, including election polling
- Especially problematic over the past ~10 years with suspicious of **non-ignorable** nonresponse:
  - Differential partisan nonresponse
    - Supporters of a particular candidate or party disproportionately not responding to the polls
    - Supporters of a particular candidate or party more enthusiastic for answering polls
  - Differences persist even after adjusting for observed demographic and other characteristics

When you have Big Problems…

…you have to bring the Big Guns!

When you have Big Problems…

…you have to bring the Big Guns!

REALLY BIG GUNS!

Nonresponse weighting adjustment

Table 1

Effect of Weighting Adjustments on Bias and Variance of a Mean, by Strength of Association of the Adjustment Cell Variables with Nonresponse and Outcome

		Association with outcome	
Association with nonresponse		Low	High
		Cell 1	Cell 3
	Low	Bias: ---	Bias: ---
		Var: ---	Var: ↓
		Cell 2	Cell 4
	High	Bias: ---	Bias: ↓
		Var: ↑	Var: ↓

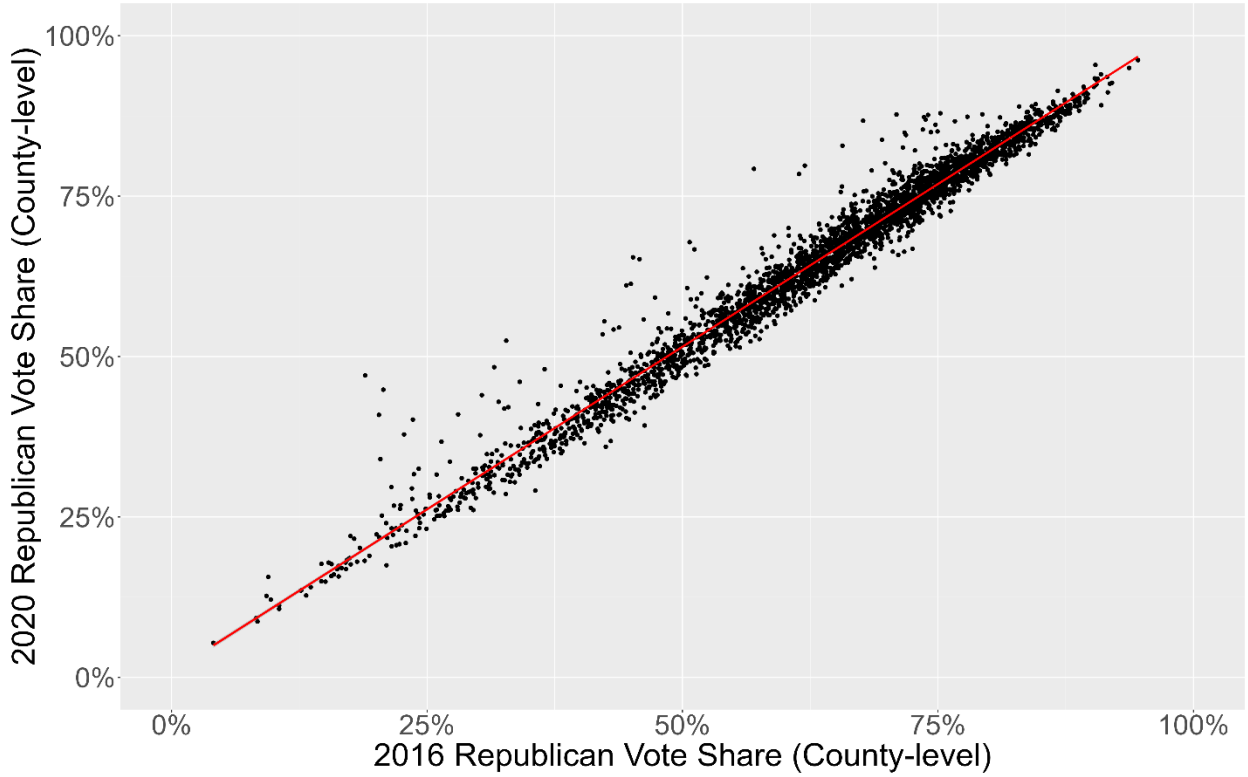
Little and Vartivarian (2005)

The Good Place

Weighting by past vote

- Past vote is the ideal variable for weighting to adjust for nonresponse in election polls:
  - Correlated with the survey outcome (voting intentions/behavior)
  - Correlated with nonresponse

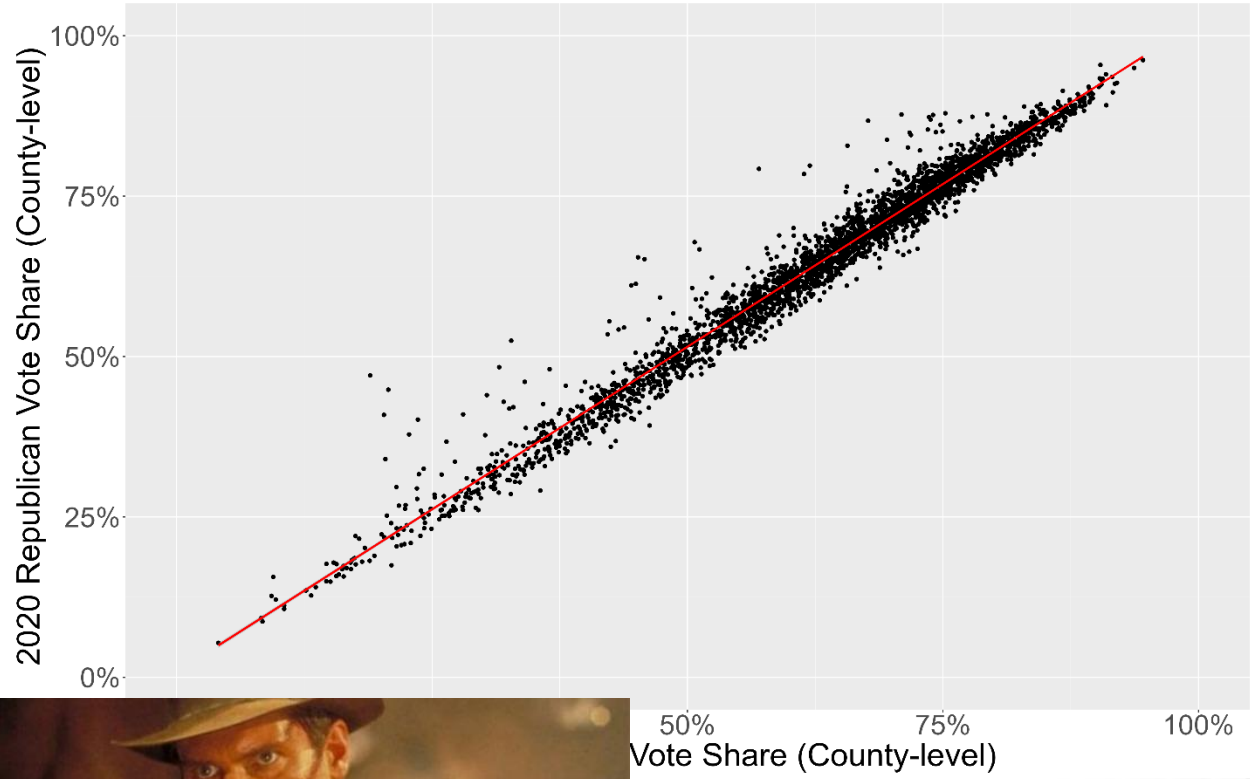
- [Survey statisticians would do anything to find a variable this powerful in other types of surveys]



Weighting by past vote

- Past vote is the ideal variable for weighting to adjust for nonresponse in election polls:
  - Correlated with the survey outcome (voting intentions/behavior)
  - Correlated with nonresponse

- [Survey statisticians would do anything to find a variable this powerful in other types of surveys]



What's the problem then?

How One Polling Decision Is Leading to Two Distinct Stories of the Election

A methodological choice has created divergent paths of polling results. Is this election more like 2020 or 2022?

Why this kind of weighting is controversial

When I started following polling methodology debates 20 years ago, weighting on recalled vote was considered a very bad idea. A surprising number of respondents don't remember how they voted; they seem likelier to remember voting for the winner; and they sometimes report voting when voting records show they did not.

10/14/2024 · 35 MIN

Weighting For Election Day

FiveThirtyEight Politics

▶ Play

With three weeks until Election Day, the 538 crew analyzes the state of what remains an extremely close presidential race. They also wade into the (very nerdy) debate over "weighting by recalled vote" that is roiling the polling community.

Learn more about your ad choices. Visit podcastchoices.com/adchoices

Weighting by recalled vote: "Bad use of polling"

Measurement error concerns in recalled vote

- Turnout overreporting (Silver et al 1986; Bernstein et al 2001; Belli et al 2001)
  - Can be reduced with voter file validation
- Recall error
- Winner's bias

Reported vote in 2012 interview	Percent Recalling Same Vote in 2016 as Reported in 2012	
	Internet %	Phone %
Barack Obama	97.6	96.9
Mitt Romney	96.1	95.2
Other/non-voter	70.0	66.7

YouGov | yougov.com October 2016

Rivers and Lauderdale (2016)

Reported 2020 vote in ANES 2024	Reported 2020 vote in ANES 2020 (col %)		
	Did not vote	Joe Biden	Donald Trump
Did not vote	75%	4%	5%
Joe Biden	11%	95%	3%
Donald Trump	15%	1%	92%

American National Election Studies (2025)

- [What about measurement error in other weighting variables?]

Research questions

1. What are the implications of measurement error in recalled vote in election polling estimates?
2. Does measurement error in recalled vote offset the expected nonresponse error reductions in weighting by it?
   - If so, what's the threshold?

Methods: Simulation study - Main variables

- Simulated a finite population of N = 1,000,000
- Vote in the current election (Y = NV, D_2, R_2) as function of:
  - Vote in the last election (X = NV, D_1, R_1)
  - Proxy covariate ($Z = Z_1, Z_2, Z_3$) for standard weighting variables (such as age, gender, education, region, etc.)

	Past in the last election (*X*) (col %)			
Current election vote (*Y*)	NV	D_1	R_1	
NV	49%	4%	6%	20%
D_2	15%	92%	4%	39%
R_2	36%	4%	89%	42%
	34%	35%	32%	

$\varphi_c = 0.68$

	Proxy covariate (*Z*) (col %)			
Current election vote (*Y*)	Z_1	Z_2	Z_3	
NV	9%	18%	33%	20%
D_2	64%	11%	39%	39%
R_2	28%	71%	29%	42%
	35%	32%	34%	

$\varphi_c = 0.36$

Methods: Simulation study - Nonresponse

- Response propensities (p) as a function of X (past vote) and Z:

$$\text{logit}(p) = -3 - 0.25NV - 0.5R_1 - 0.1Z_2 - 0.2Z_3$$

- Overall Response Rate ≈ 3.5%
- Response Rate by group:

X	RR
NV	3.4%
D_1	4.4%
R_1	2.7%

Z	RR
Z_1	4.1%
Z_2	3.1%
Z_3	3.2%

Methods: Simulation study – Measurement error

- Simulated six types of measurement error in recalled vote
  - Turnout overreporting
    - X_1: Disproportionately to D_1
    - X_2: Disproportionately to R_1
  - Vote misreport $R_1 \rightarrow D_1$
    - X_3: moderate
    - X_4: strong
  - Vote misreport $D_1 \rightarrow R_1$
    - X_5: moderate
    - X_6: strong

Methods: Overreporting of turnout

Turnout overreporting disproportionately to D_1

	Actual Past Vote (*X*) (col %)			
Recalled Vote (X_1)	NV	D_1	R_1	
NV	42%	1%	1%	15%
D_1	32%	98%	1%	45%
R_1	26%	1%	98%	40%
	34%	35%	32%	

$\varphi_c = 0.72$

Turnout overreporting disproportionately to R_1

	Actual Past Vote (*X*) (col %)			
Recalled Vote (X_2)	NV	D_1	R_1	
NV	42%	1%	1%	15%
D_1	26%	98%	1%	43%
R_1	32%	1%	98%	42%
	34%	35%	32%	

$\varphi_c = 0.72$

Methods: Vote misreport from $R_1 \rightarrow D_1$

Moderate vote misreport from $R_1 \rightarrow D_1$

	Actual Past Vote (*X*) (col %)			
Recalled Vote (X_3)	NV	D_1	R_1	
NV	98%	2%	2%	34%
D_1	1%	98%	10%	38%
R_1	1%	1%	87%	28%
	34%	35%	32%	

$\varphi_c = 0.92$

Strong vote misreport from $R_1 \rightarrow D_1$

	Actual Past Vote (*X*) (col %)			
Recalled Vote (X_4)	NV	D_1	R_1	
NV	98%	2%	2%	34%
D_1	1%	98%	40%	47%
R_1	1%	1%	58%	19%
	34%	35%	32%	

$\varphi_c = 0.81$

Methods: Vote misreport from $D_1 \rightarrow R_1$

Moderate vote misreport from $D_1 \rightarrow R_1$

	Actual Past Vote (*X*) (col %)			
Recalled Vote (X_5)	NV	D_1	R_1	
NV	98%	3%	1%	34%
D_1	1%	87%	1%	31%
R_1	1%	10%	98%	35%
	34%	35%	32%	

$\varphi_c = 0.91$

Strong vote misreport from $D_1 \rightarrow R_1$

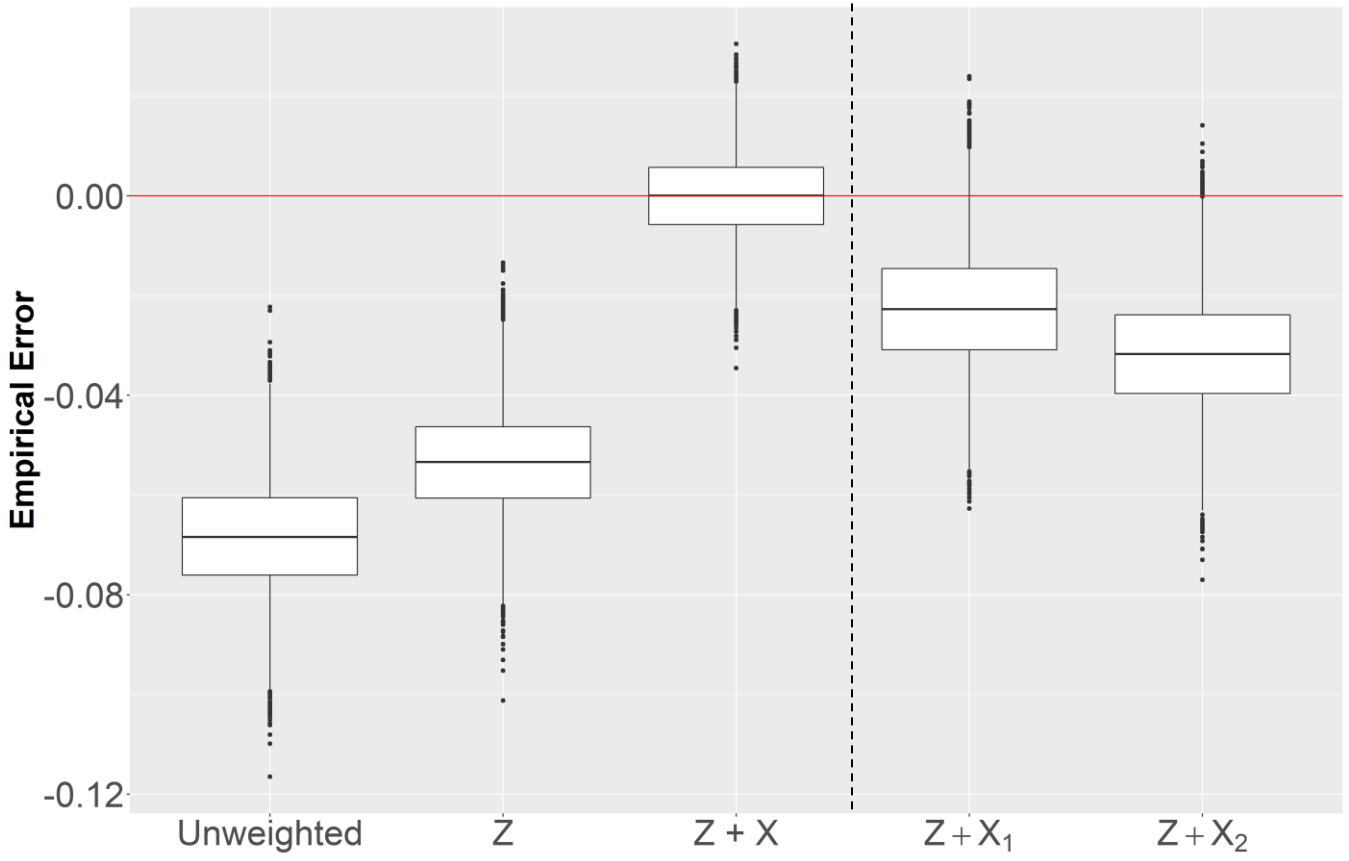
	Actual Past Vote (*X*) (col %)			
Recalled Vote (X_6)	NV	D_1	R_1	
NV	98%	2%	2%	34%
D_1	1%	58%	1%	21%
R_1	1%	40%	97%	45%
	34%	35%	32%	

$\varphi_c = 0.80$

Methods: Simulation study – Weighting adjustments

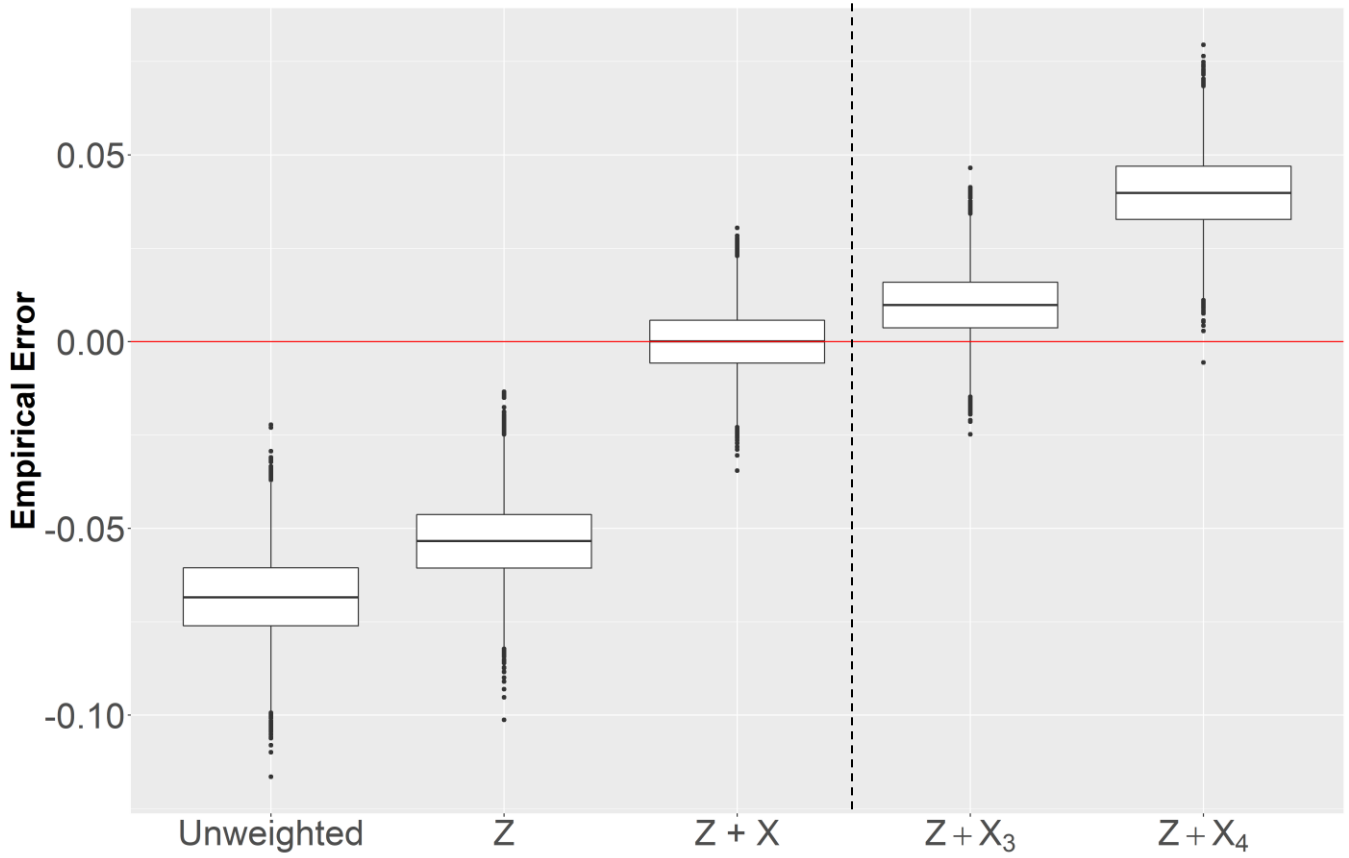
- Calibrated the respondent sample by:
  - Proxy covariate Z only (Z)
  - Past vote X only (X)
  - Both Z and X ($Z + X$)
  - Both Z and recalled vote with measurement error, $X_1 - X_6$ ($Z + X_i$)
- Simulated K = 10,000 simple random samples of size n = 50,000, leading up to approximately 1,750 respondents
- Assessed the properties of the proportion of support to candidate R_2 of each estimator case across all simulation runs by:
  - Empirical bias
  - Empirical standard error
  - Empirical root mean squared error

Results: Turnout overreporting



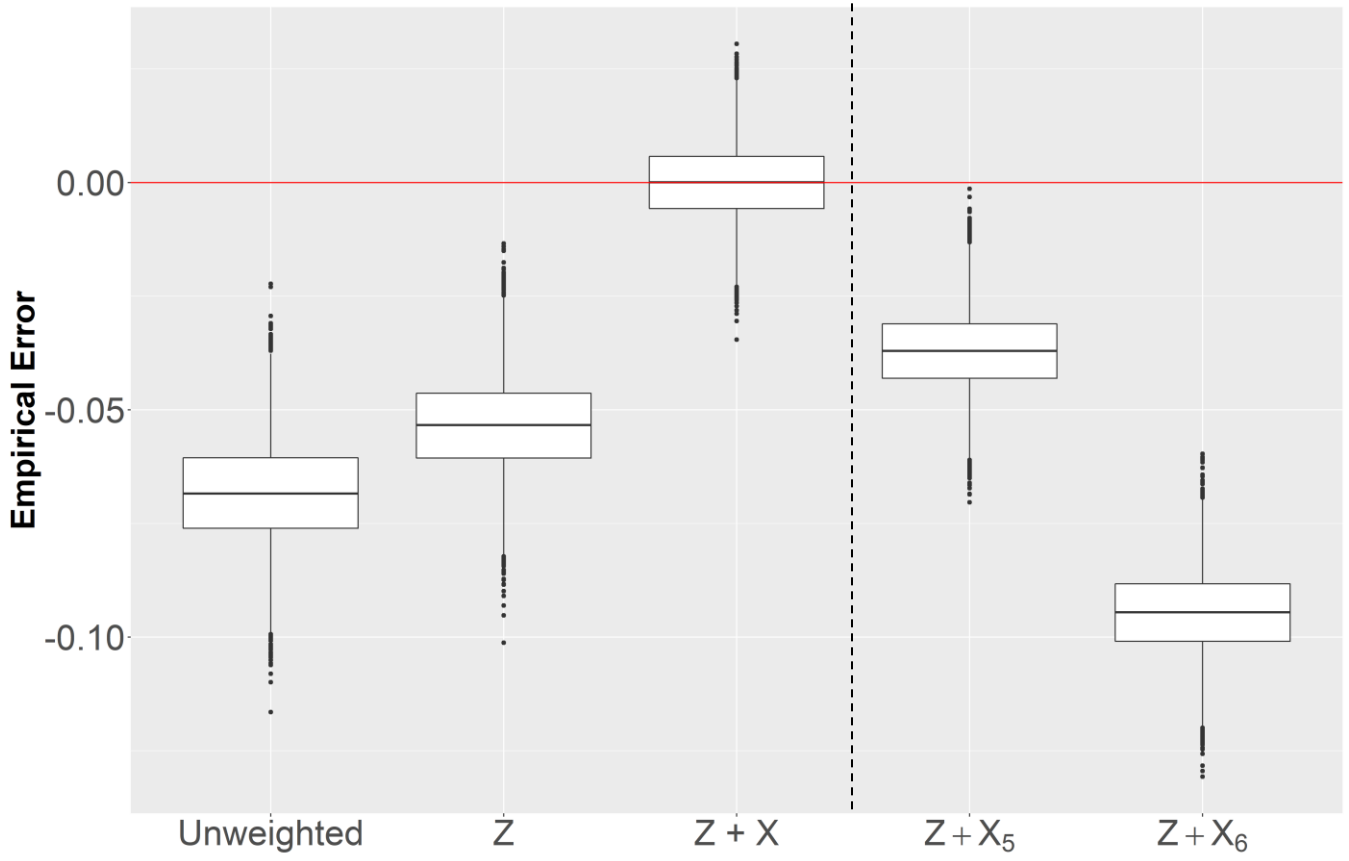
Emp Bias (x 100)	-6.834	-5.339	0.002	-2.258	-3.173
Emp Std Error (x 100)	1.139	1.066	0.844	1.190	1.174
Emp RMSE (x 100)	6.834	5.339	0.002	2.258	3.173

Results: Vote misreport from $R_1 \rightarrow D_1$



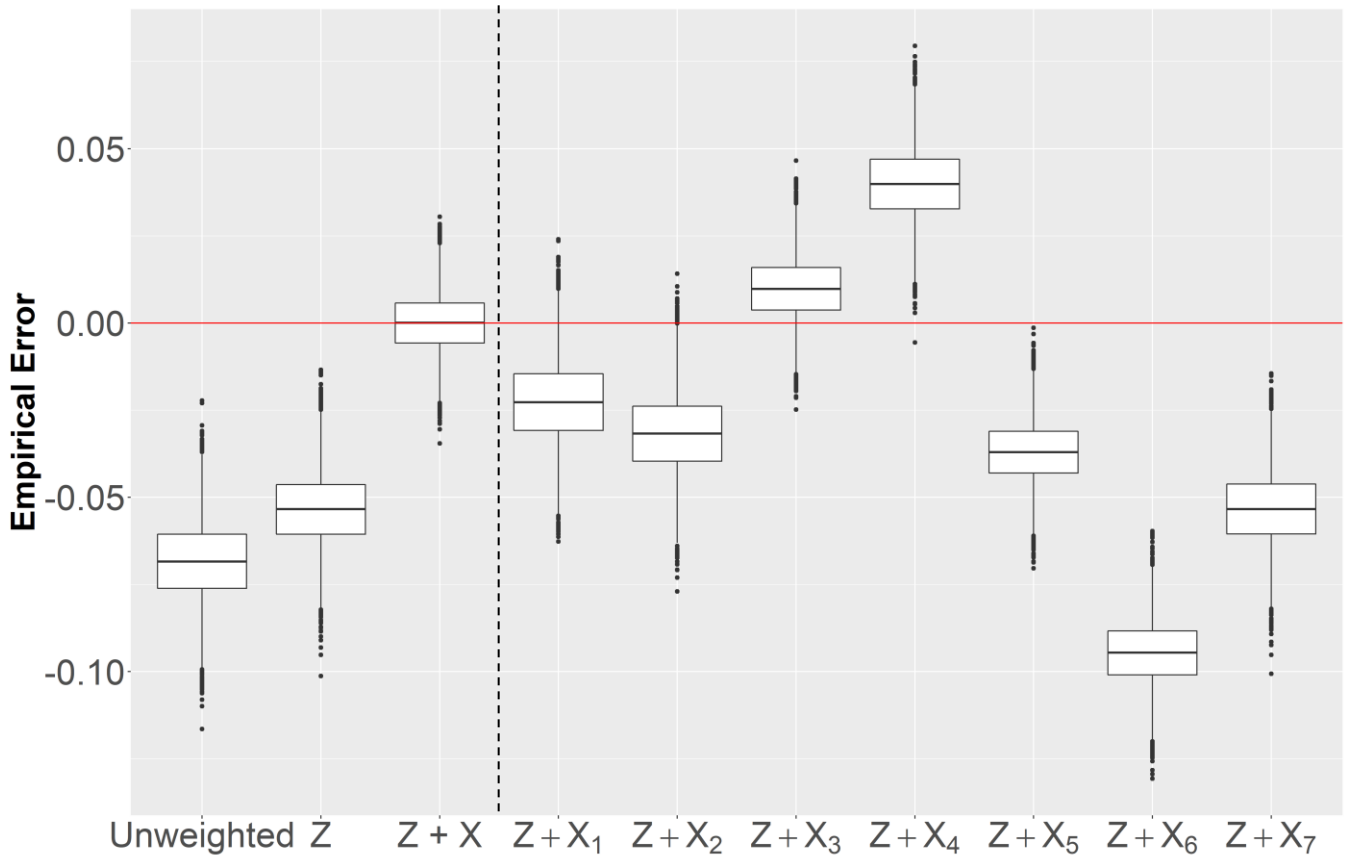
Emp Bias (x 100)	-6.834	-5.339	0.002	0.980	3.982
Emp Std Error (x 100)	1.139	1.066	0.844	0.909	1.038
Emp RMSE (x 100)	6.834	5.339	0.002	0.980	3.982

Results: Vote misreport from $D_1 \rightarrow R_1$



Emp Bias (x 100)	-6.834	-5.339	0.002	-3.707	-9.460
Emp Std Error (x 100)	1.139	1.066	0.844	0.905	0.941
Emp RMSE (x 100)	6.834	5.339	0.002	3.707	9.460

Results: Summary



Emp Bias (x 100)	-6.834	-5.339	0.002	-2.258	-3.173	0.980	3.982	-3.707	-9.460	-5.332
Emp Std Error (x 100)	1.139	1.066	0.844	1.190	1.174	0.909	1.038	0.905	0.941	1.058
Emp RMSE (x 100)	6.834	5.339	0.002	2.258	3.173	0.980	3.982	3.707	9.460	5.332

Conclusions

- In most cases, the measurement error in the recall vote did not offset the gains in accuracy due to weighting on it
  - Unless the measurement error was substantial with voting misreporting taking place disproportionately more towards the support to the losing candidate AND their voters tend to be less likely to vote
  - In this simulation, the threshold was about 15% reporting they voted for R_1 among those who voted for D_1
- Overall, better off weighting by recall vote than not

Limitations and Next Steps

- Limitations:
  - Results conditioned to simulation parameters, including:
    - Relationship between vote in the past and current elections
    - Relationship between past vote and nonresponse
    - Margins between the two candidate in the past and current elections
  - Only look at trade-off between measurement error and nonresponse
- Next steps:
  - Assess other simulation scenarios
  - Evaluate these results when nonresponse is non-ignorable

Discussion

- Recalled vote is a powerful variable for nonresponse weighting adjustments of election polls/surveys
- Weighting by recalled vote is not a panacea:
  - Better to deal with nonresponse at the design/data collection stage
  - Does not removed bias due to extreme cases of non-ignorable nonresponse
    - Nonresponse directly associated to vote in the current election
- Challenges with recalled vote weighting:
  - Statewide polls: movers
  - Voter file matching/validation
  - Different was to implement it!

References

- American National Election Studies (2025). ANES 2024 Time Series Study Preliminary Release: Combined Pre-Election and Post-Election Data [dataset and documentation]. April 30, 2025 version. www.electionstudies.org
- Belli, R. F., Traugott, M. W., & Beckmann, M. N. (2001). What leads to voting overreports? Contrasts of overreporters to validated voters and admitted nonvoters in the American National Election Studies. *Journal of Official Statistics*, 17(4), 479.
- Bernstein, R., Chadha, A., & Montjoy, R. (2001). Overreporting voting: Why it happens and why it matters. *Public Opinion Quarterly*, 65(1), 22-44.
- Little, R. J., & Vartivarian, S. (2005). Does weighting for nonresponse increase the variance of survey means? *Survey Methodology*, 31(2), 161.
- Rivers, D., Lauderdale, B. (2016). Beware the phantom swings: why dramatic bounces in the polls aren't always what they seem. https://today.yougov.com/politics/articles/16873-beware-phantom-swings-why-dramatic-swings-in-the-p
- Silver, B. D., Anderson, B. A., & Abramson, P. R. (1986). Who overreports voting? *American Political Science Review*, 80(2), 613-624.

INSTITUTE FOR SOCIAL RESEARCH • SURVEY RESEARCH CENTER
SURVEY RESEARCH OPERATIONS
UNIVERSITY OF MICHIGAN

Thank you!

raphaeln@umich.edu