

V0D 2018 CTC 1

Why 25% of Voter Polls are Wrong

by
Milo Schield, Augsburg University
Fellow: American Statistical Association
US Rep: International Statistical Literacy Project
 December 2, 2018
Twin Cities Critical Thinking Club (CTC)
 Slides: www.StatLit.org/pdf/2018-Schield-CTC-Slides.pdf

V0D 2018 CTC 2

25% of Voter Polls are Wrong!

Polls are short surveys.

- Less demographic information.
- Fewer choices for answers.

Polls typically involve sampling: sampling error.
 Voter polls are of wide-spread interest.

25%* of voter polls in the final 21 days are wrong.
 Wrong: actual result outside 95% confidence interval
 In 2016, 29%** of those in final 21 days were wrong.
* <https://fivethirtyeight.com/features/how-the-fivethirtyeight-senate-forecast-model-works/>
 ** <https://fivethirtyeight.com/features/the-polls-are-all-right/>

V0D 2018 CTC 3

Random Sampling: Non-Response Bias

Surveys and polls often interview people by phone.
 Plus: Cheap; random sample.
 Minus: Non-response bias.
 People who respond are not representative of the entire population.

Year	Response Rate (%)
1997	36
2000	28
2003	25
2006	21
2009	15
2012	9
2016	9

The lower the response rate, the greater the risk of bias.

http://www.pewresearch.org/2017/05/15/what-low-response-rates-mean-for-telephone-surveys/pw-05-15-2017_rdnnonresponse-00-14/

V0D 2018 CTC 4

Voter Polls: Before or After

For a given election, voting polls are either

- before the election (opinion polls) or
- after the election (exit polls).

- Opinion polls *forecast*: use hypothetical questions
- Exit polls *explain*: use factual questions

Explanatory polls just *tabulate* – like sports statistics
 Forecast polls use *models* – like weather forecasts

V0D 2018 CTC 5

Chance of Winning: "It depends"

Election (forecast) polls are harder to interpret!

Consider two candidates in different contests
 Both have 52% of their respective votes.

- Candidate #1 has 60% chance of winning.
- Candidate #2 has a 75% chance of winning.

Q. How can they have different chances of winning?
 A. It depends on the size of the **polling error!**

V0D 2018 CTC 6

Error in Forecast Polls: Sampling and Forecast

Forecast polls have sampling error (theoretical):

- Difference b/t sample and population statistic
- Entirely due to random sampling
- Decreases as sample size increases

Forecast polls also have modeling error (empirical)

- Difference between the forecast and actual result
- Due to different models of undecided
- Independent of sample size

VOD 2018 CTC 7

Average Opinion Polling Error 1972-2016

Popular Vote: Polls' error in final 21 days

Cycle	National	State	Cycle	National	State
2016	3.1	5.2	1992	4.6	5.2
2012	3.3	3.7	1988	3.5	5
2008	2.3	3.9	1984	5.4	4.5
2004	2.2	3.5	1980	8.9	8.6
2000	3.9	4.6	1976	2.5	3.8
1996	6.4	4.8	1972	2.6	4.6
Ave	3.5	4.3	Ave	4.6	5.3

<https://fivethirtyeight.com/features/the-polls-are-all-right/>

VOD 2018 CTC 8

Opinion Polling Error Bigger than Sampling Error

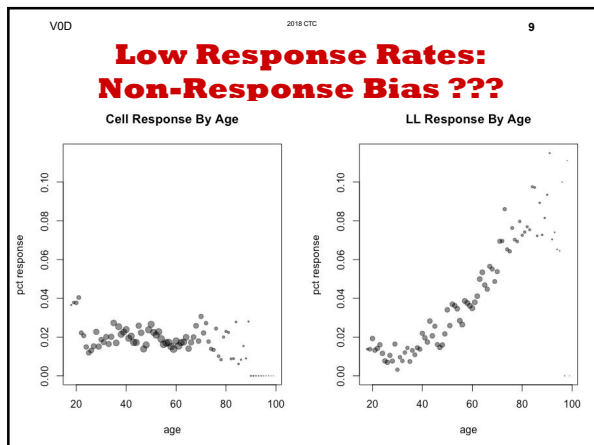
Average popular-vote polling error* in final 21 days:

- 4 points: Presidential
- 5.4 points: US Senate and Governor
- 6.2 points: US House
- 8.9 points: Primary

* D-R margin error. <https://fivethirtyeight.com/features/the-polls-are-all-right/>

95% sampling error by sample size:

- 3.3 points: n = 900
- 3.0 points: n = 1,024 [Most common sample size]
- 2.8 points: n = 1,225

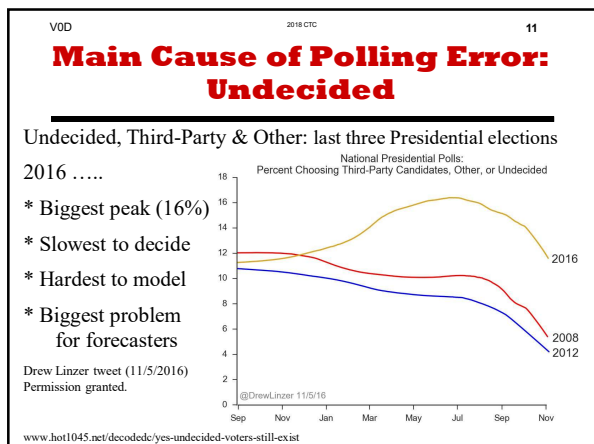


VOD 2018 CTC 10

Causes of Error in Opinion Polls

Causes of error unique to opinion (forecast) polls:

1. Respondents **undecided** about whether to vote
2. Respondents **undecided** about who to vote for
3. Respondents change their mind before voting
4. Voter opinion polls must forecast
 - which **undecided** will vote
 - who **undecided** voters will vote for
 - which **third-party voters** will change their mind
5. Different polls allocate undecided differently
6. Polls allocate using demographic/historical data



VOD 2018 CTC 12

Conclusions

To interpret polls, know the difference:

1. between an opinion poll, an exit poll and a survey
2. between polling error and sampling error

To evaluate polls, know that

1. the chance of winning depends on polling error
2. polling error is bigger than sampling error
3. polling error exists because opinion polls are forecasts
4. the size of polling error depends on the contest

Q. How can pollsters educate readers on polling error?

Why 25% of Voter Polls are Wrong

by

Milo Schield, Augsburg University

Fellow: American Statistical Association

US Rep: International Statistical Literacy Project

December 2, 2018

Twin Cities Critical Thinking Club (CTC)

Slides: www.StatLit.org/pdf/2018-Schild-CTC-Slides.pdf

25% of Voter Polls are Wrong!

Polls are short surveys.

- Less demographic information.
- Fewer choices for answers.

Polls typically involve sampling: sampling error.

Voter polls are of wide-spread interest.

25%* of voter polls in the final 21 days are wrong.

Wrong: actual result outside 95% confidence interval

In 2016, 29%** of those in final 21 days were wrong.

* <https://fivethirtyeight.com/features/how-the-fivethirtyeight-senate-forecast-model-works/>

** <https://fivethirtyeight.com/features/the-polls-are-all-right/>

Random Sampling: Non-Response Bias

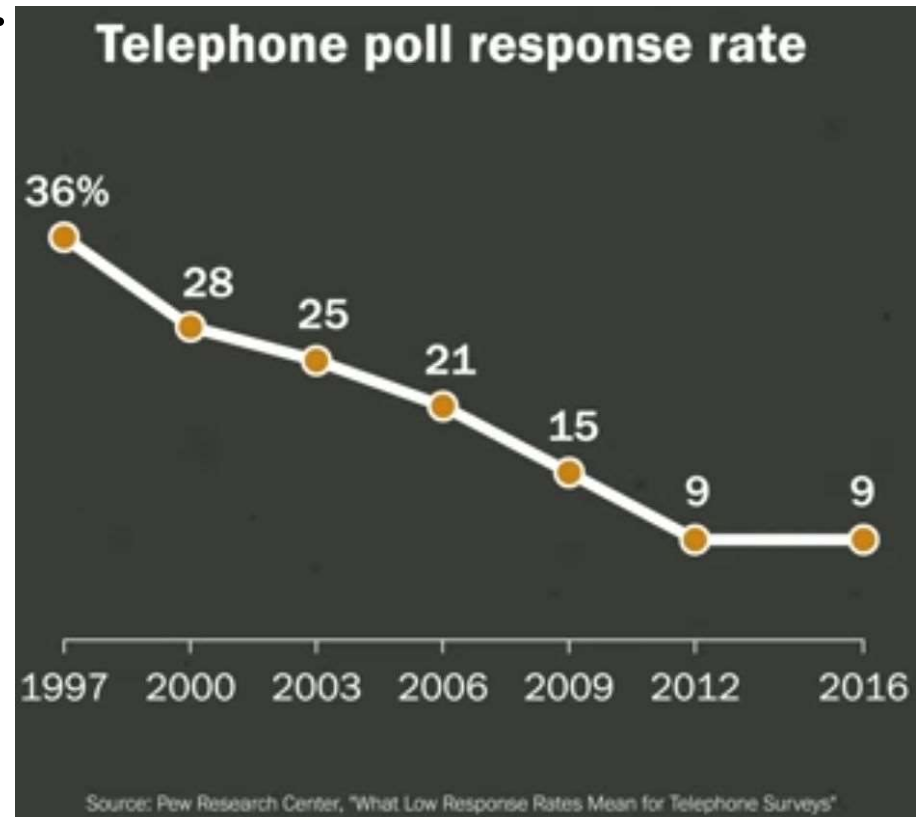
Surveys and polls often interview people by phone.

Plus: Cheap; random sample.

Minus: Non-response bias.

People who respond are
not representative of the
entire population.

The lower the response rate,
the greater the risk of bias.



Voter Polls: Before or After

For a given election, voting polls are either

- before the election (opinion polls) or
- after the election (exit polls).

- Opinion polls *forecast*: use hypothetical questions
- Exit polls *explain*: use factual questions

Explanatory polls just *tabulate* – like sports statistics

Forecast polls use *models* – like weather forecasts

Chance of Winning: “It depends”

Election (forecast) polls are harder to interpret!

Consider two candidates in different contests

Both have 52% of their respective votes.

- Candidate #1 has 60% chance of winning.
- Candidate #2 has a 75% chance of winning.

Q. How can they have different chances of winning?

A. It depends on the size of the **polling error!**

Error in Forecast Polls: Sampling and Forecast

Forecast polls have sampling error (theoretical):

- Difference b/t sample and population statistic
- Entirely due to random sampling
- Decreases as sample size increases

Forecast polls also have modeling error (empirical)

- Difference between the forecast and actual result
- Due to different models of undecided
- Independent of sample size

Average Opinion Polling Error 1972-2016

Popular Vote: Polls' error in final 21 days						
Cycle	National	State		Cycle	National	State
2016	3.1	5.2		1992	4.6	5.2
2012	3.3	3.7		1988	3.5	5
2008	2.3	3.9		1984	5.4	4.5
2004	2.2	3.5		1980	8.9	8.6
2000	3.9	4.6		1976	2.5	3.8
1996	6.4	4.8		1972	2.6	4.6
Ave	3.5	4.3		Ave	4.6	5.3

<https://fivethirtyeight.com/features/the-polls-are-all-right/>

Opinion Polling Error Bigger than Sampling Error

Average popular-vote polling error* in final 21 days:

- 4 points: Presidential
- 5.4 points: US Senate and Governor
- 6.2 points: US House
- 8.9 points: Primary

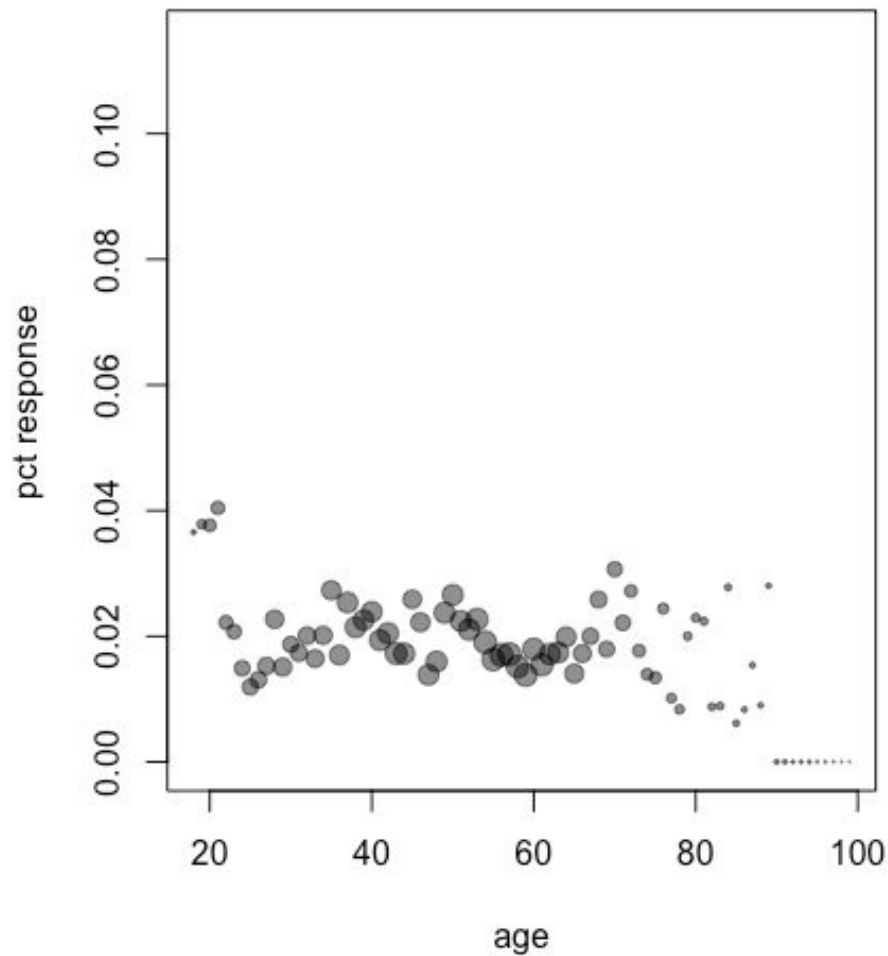
* D-R margin error. <https://fivethirtyeight.com/features/the-polls-are-all-right/>

95% sampling error by sample size:

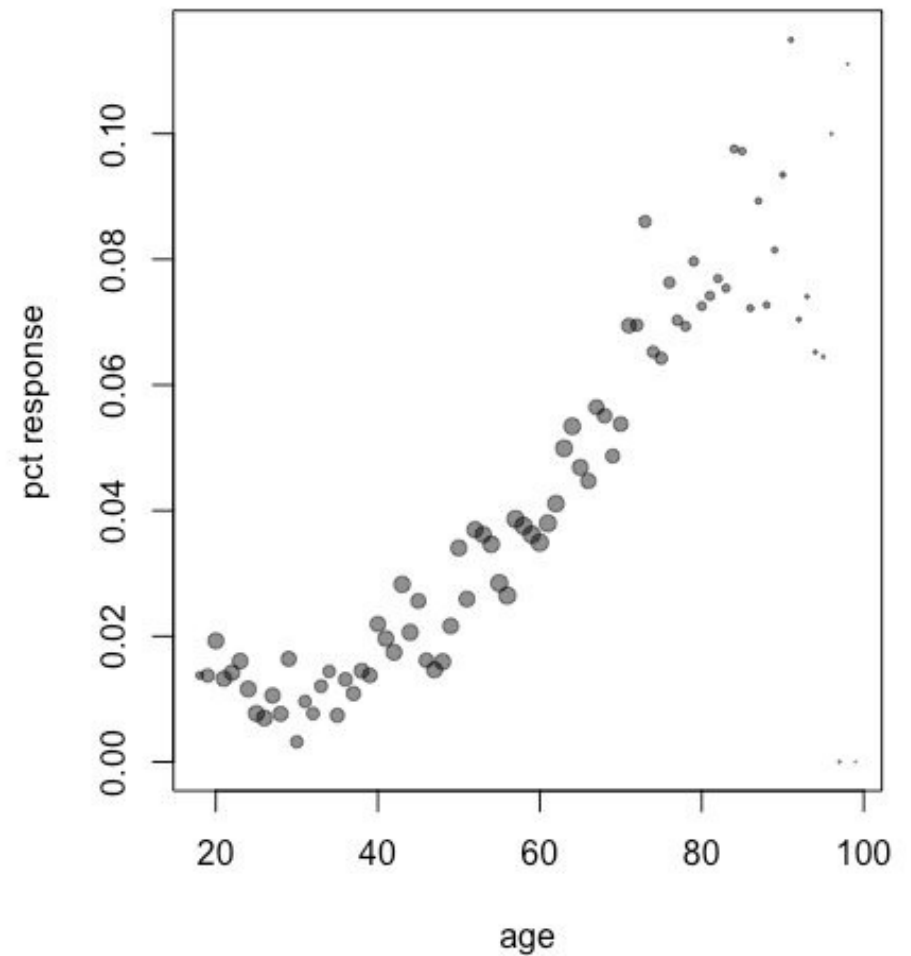
- 3.3 points: $n = 900$
- 3.0 points: $n = 1,024$ [Most common sample size]
- 2.8 points: $n = 1,225$

Low Response Rates: Non-Response Bias ???

Cell Response By Age



LL Response By Age



Causes of Error in Opinion Polls

Causes of error unique to opinion (forecast) polls:

1. Respondents **undecided** about whether to vote
2. Respondents **undecided** about who to vote for
3. Respondents change their mind before voting
4. Voter opinion polls must forecast
 - which **undecided** will vote
 - who **undecided** voters will vote for
 - which **third-party voters** will change their mind
5. Different polls allocate undecided differently
6. Polls allocate using demographic/historical data

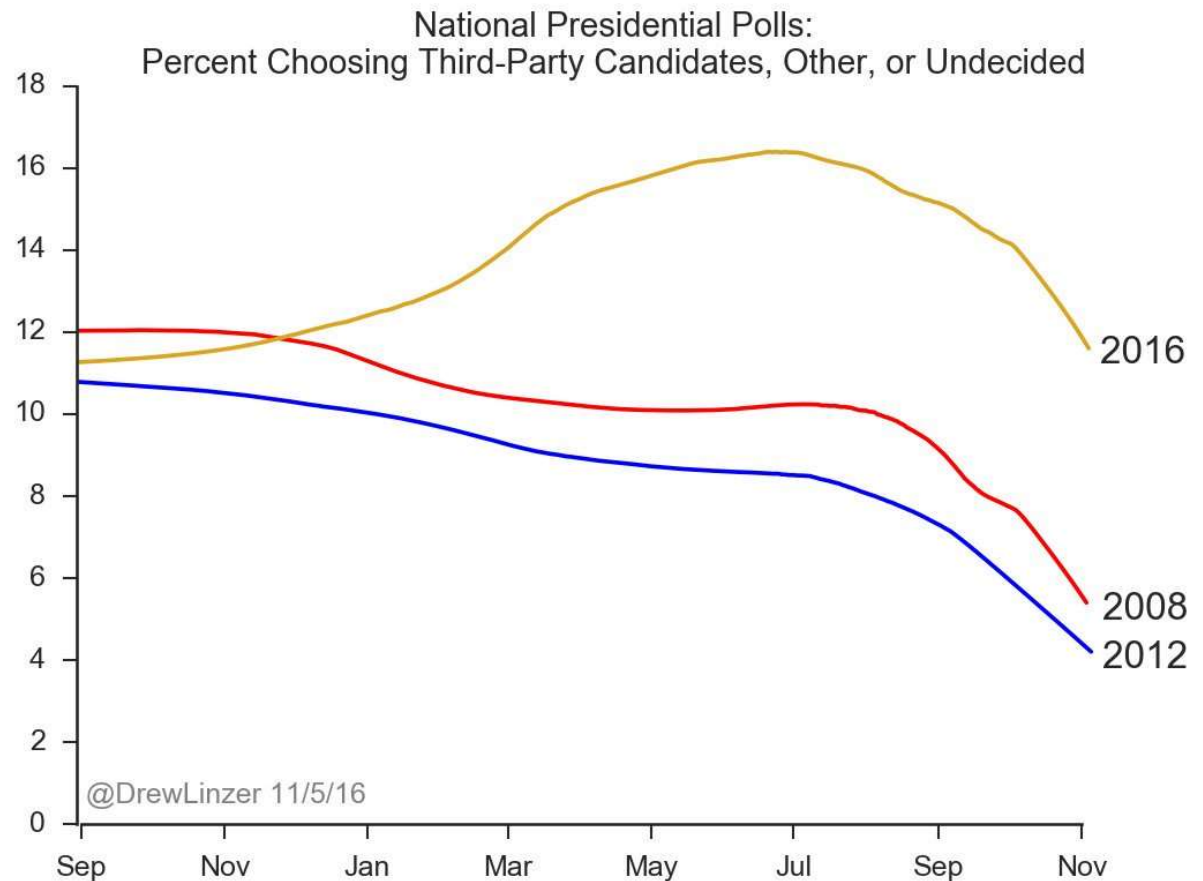
Main Cause of Polling Error: Undecided

Undecided, Third-Party & Other: last three Presidential elections

2016

- * Biggest peak (16%)
- * Slowest to decide
- * Hardest to model
- * Biggest problem for forecasters

Drew Linzer tweet (11/5/2016)
Permission granted.



Conclusions

To interpret polls, know the difference:

1. between an opinion poll, an exit poll and a survey
2. between polling error and sampling error

To evaluate polls, know that

1. the chance of winning depends on polling error
2. polling error is bigger than sampling error
3. polling error exists because opinion polls are forecasts
4. the size of polling error depends on the contest

Q. How can pollsters educate readers on polling error?