Poll Aggregation in the 2016 U.S. Presidential Election: Accuracy and Uncertainty

Drew Linzer, Ph.D.
Chief Scientist, Civiqs

December 6, 2016

PUBLIC POLLING VS. PRIVATE POLLING

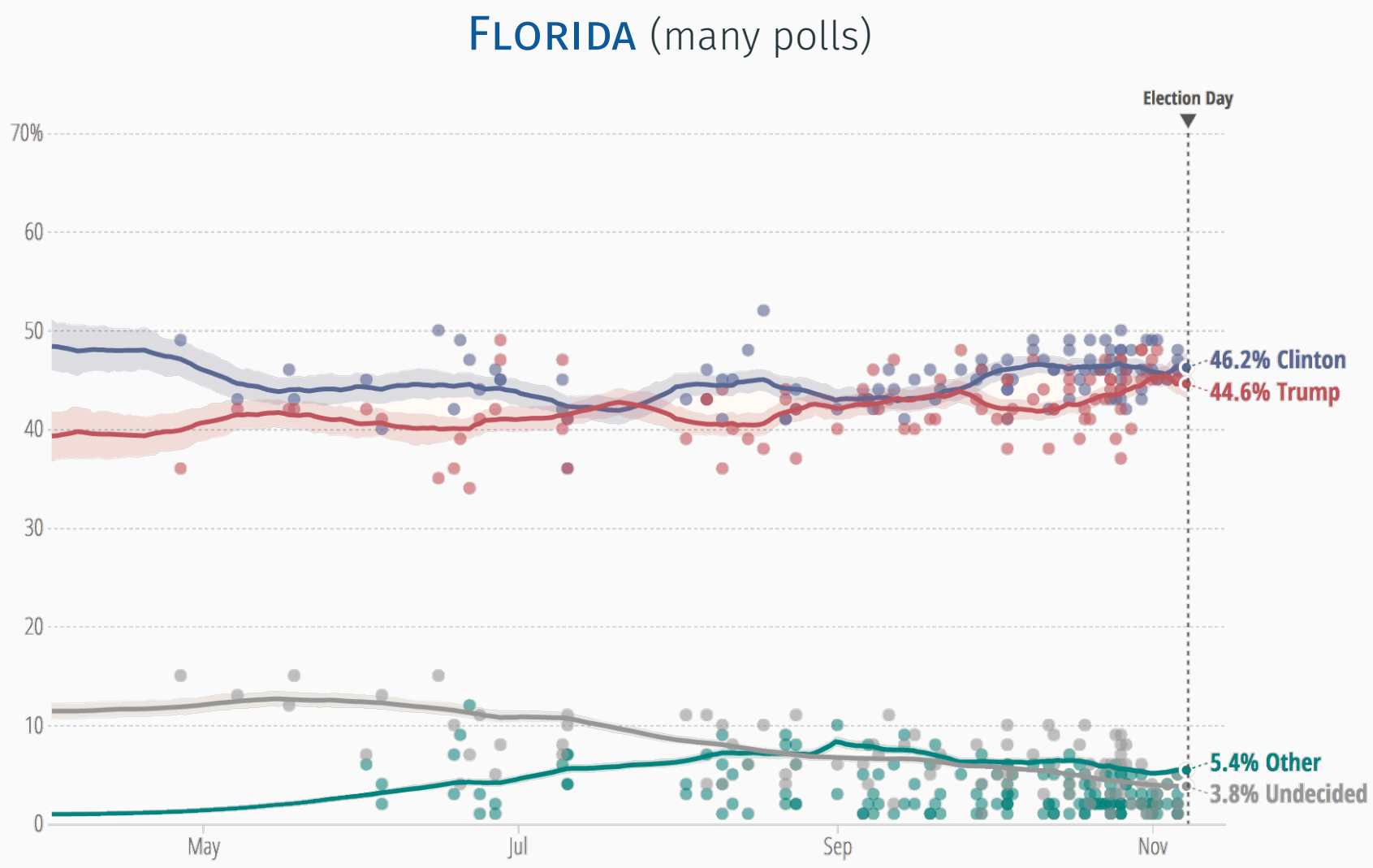
Public: 1,400 state-level presidential polls in all 50 states

- Released by media organizations and polling firms
- Quality varies, but they're **usually accurate**, on average

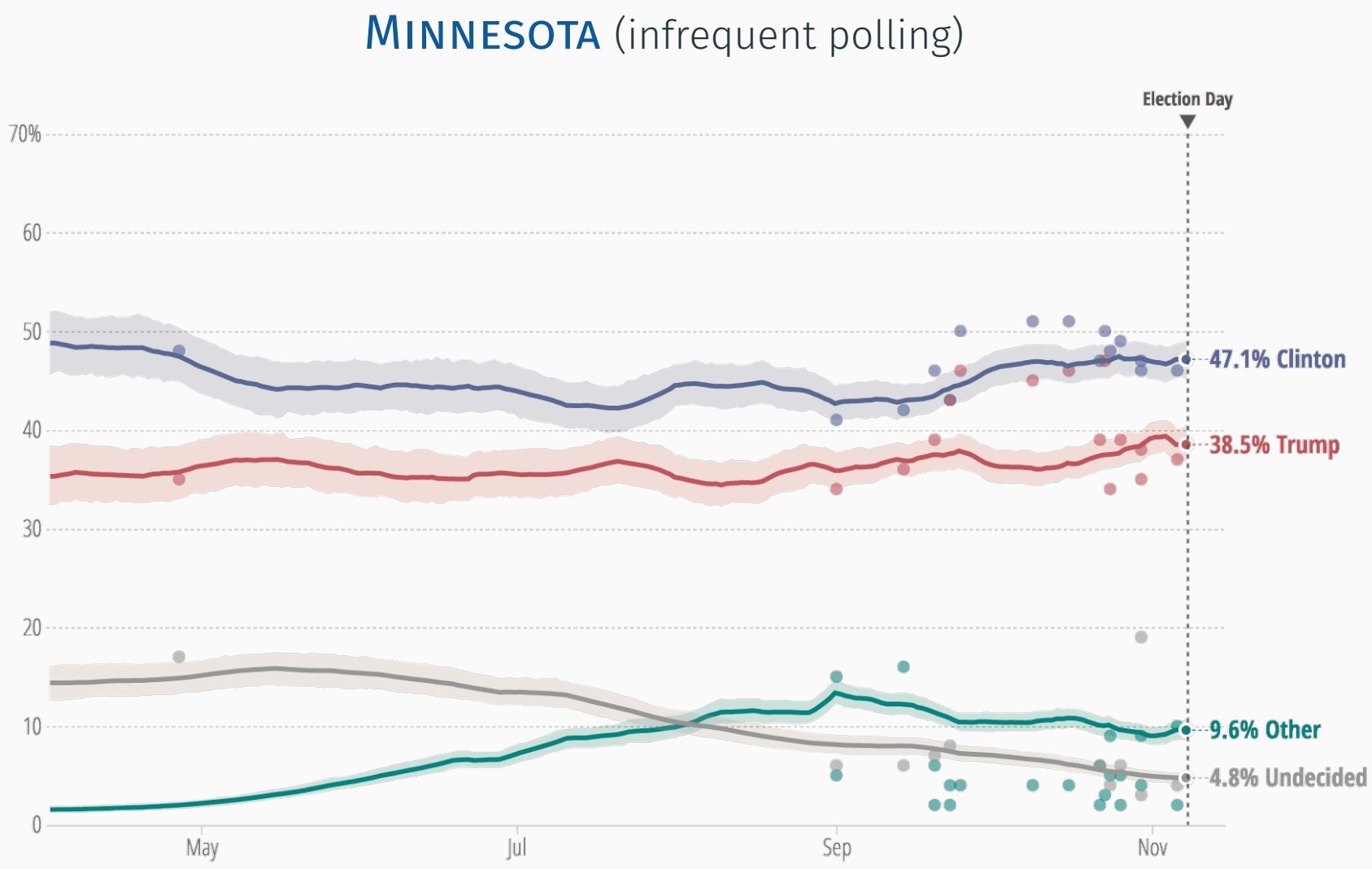
Private: Campaigns' internal research, kept secret

- Supposed to be higher quality than public polls
- Except, **neither campaign** expected Trump to win

POLL AGGREGATES CANCEL OUT RANDOM SAMPLING ERROR

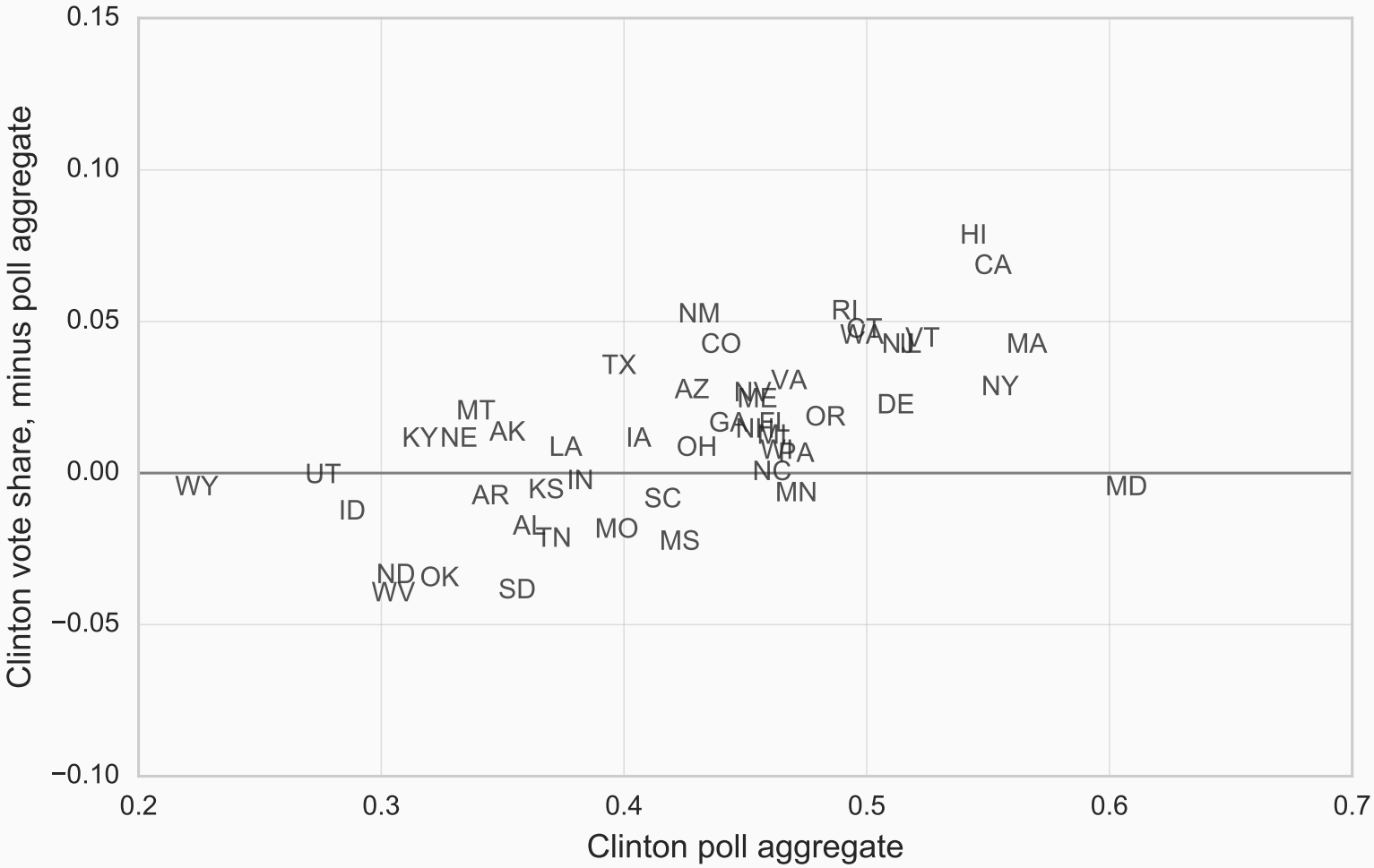


POLL AGGREGATES CANCEL OUT RANDOM SAMPLING ERROR



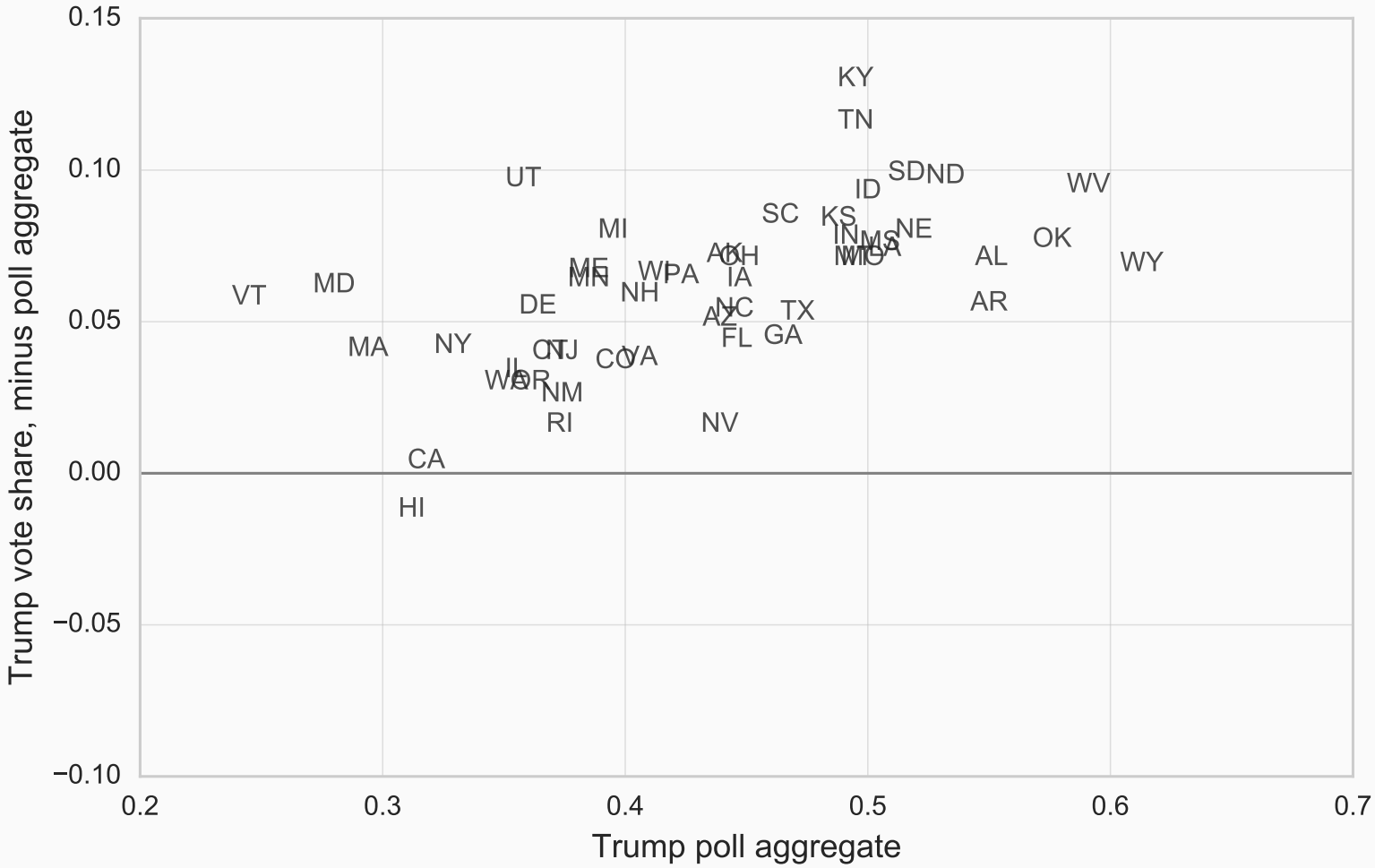
THE 2016 PRESIDENTIAL POLLS HAD SYSTEMATIC BIASES

Clinton outperformed her polls in more Democratic states



THE 2016 PRESIDENTIAL POLLS HAD SYSTEMATIC BIASES

Trump dramatically exceeded expectations from the polls



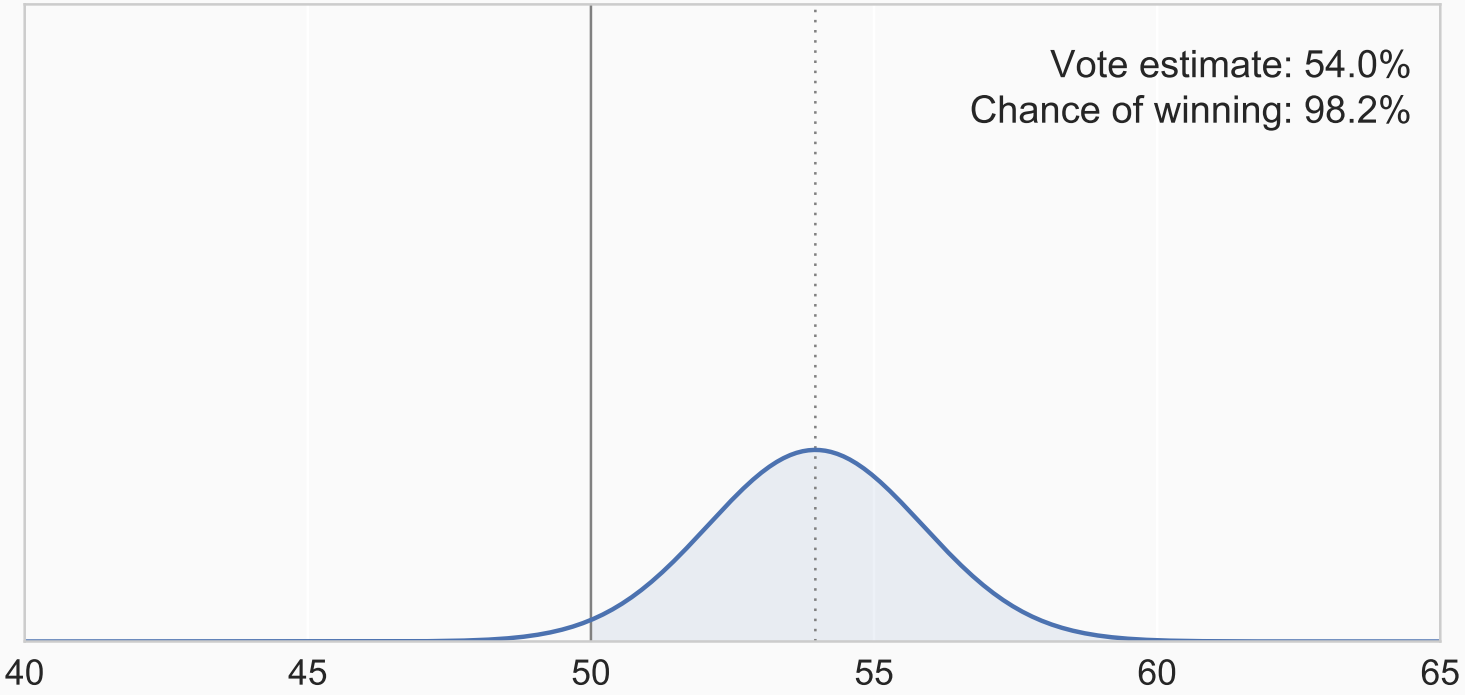
THE BIG QUESTION

How uncertain should we have been about the polls to make **5 to 10 percentage point errors** seem consistent—even minimally—with the data?

Remember: If you can't predict the bias, you have to assume the errors can go in **either direction**.

Clinton vote share, Wisconsin

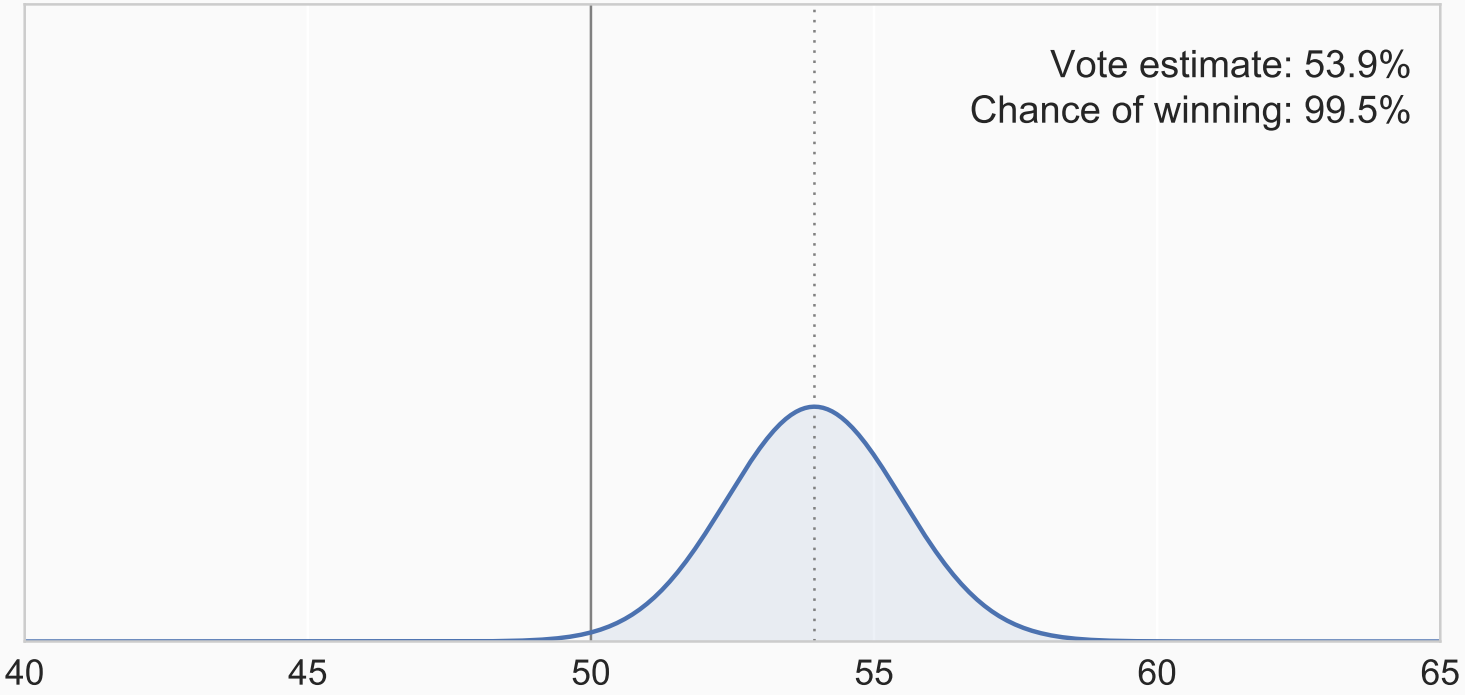
Posterior distribution, assuming simple random sampling



Date	Firm	Trump	Clinton	N
11/1/16	PPP	41%	48%	891

Clinton vote share, Wisconsin

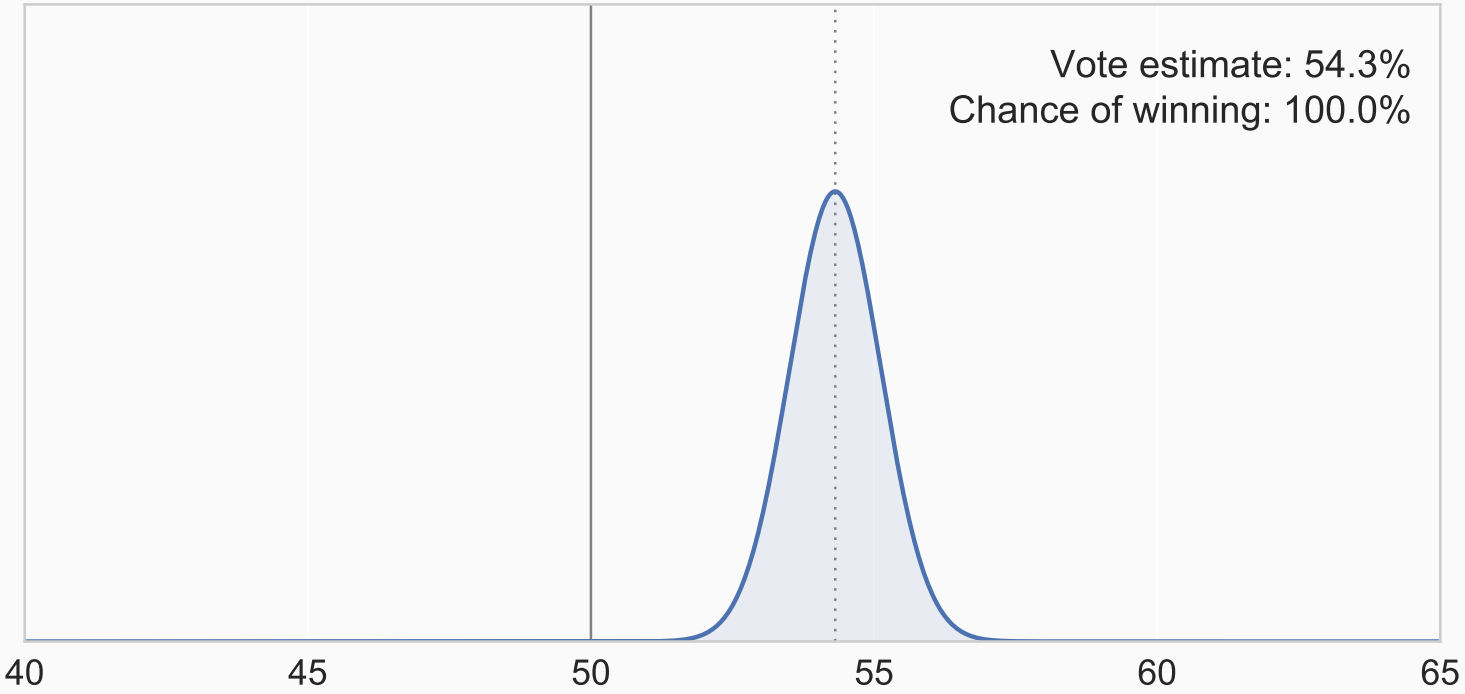
Posterior distribution, assuming simple random sampling



Date	Firm	Trump	Clinton	N
11/1/16	PPP	41%	48%	891
11/1/16	Loras College	38%	44%	500

Clinton vote share, Wisconsin

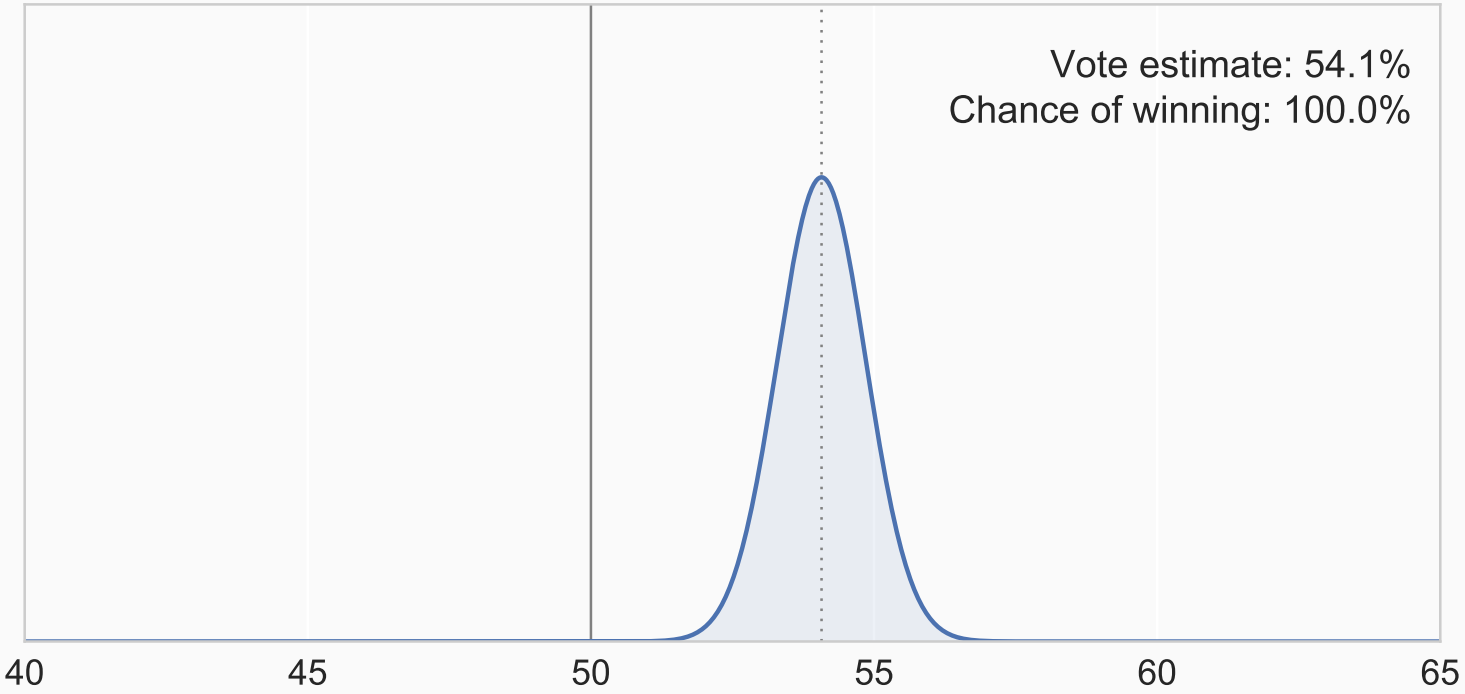
Posterior distribution, assuming simple random sampling



Date	Firm	Trump	Clinton	N
11/1/16	PPP	41%	48%	891
11/1/16	Loras College	38%	44%	500
11/2/16	Remington	41%	49%	2720

Clinton vote share, Wisconsin

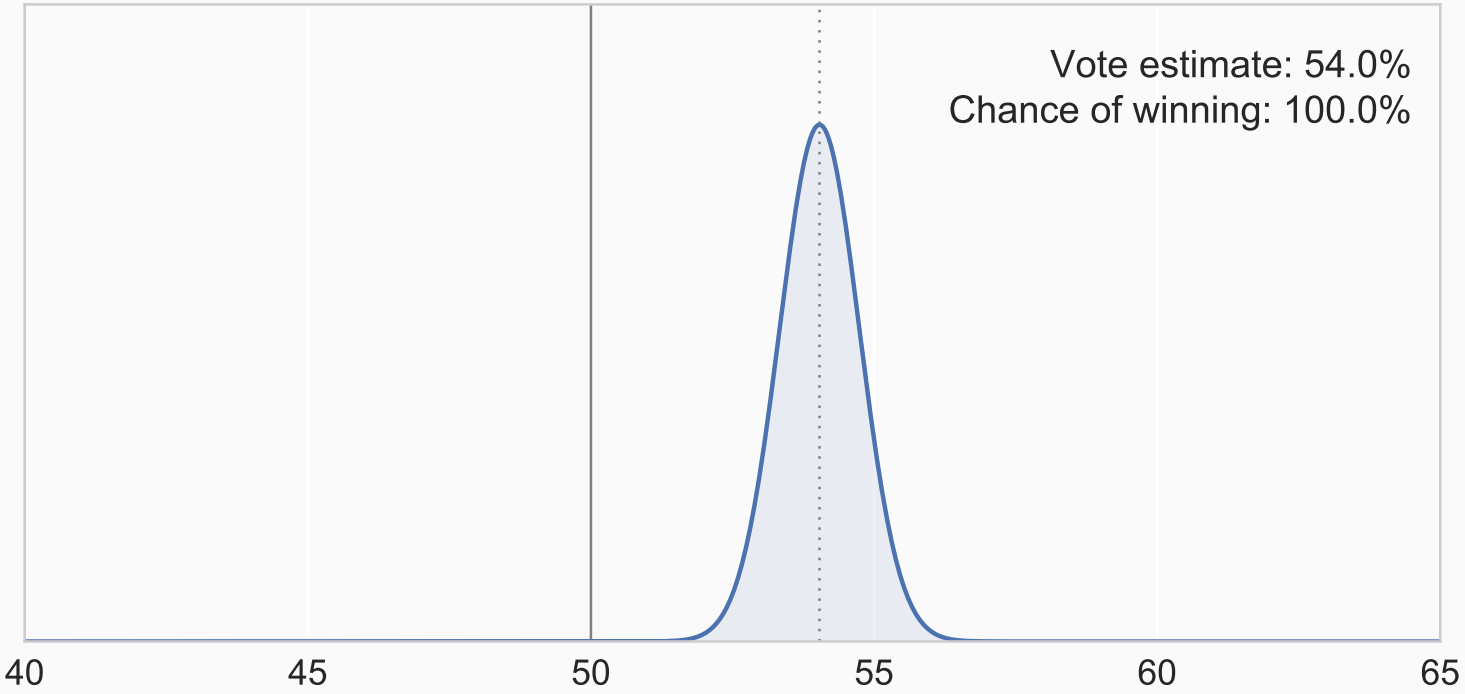
Posterior distribution, assuming simple random sampling



Date	Firm	Trump	Clinton	N
11/1/16	PPP	41%	48%	891
11/1/16	Loras College	38%	44%	500
11/2/16	Remington	41%	49%	2720
11/6/16	Ipsos/Reuters	40%	46%	842

Clinton vote share, Wisconsin

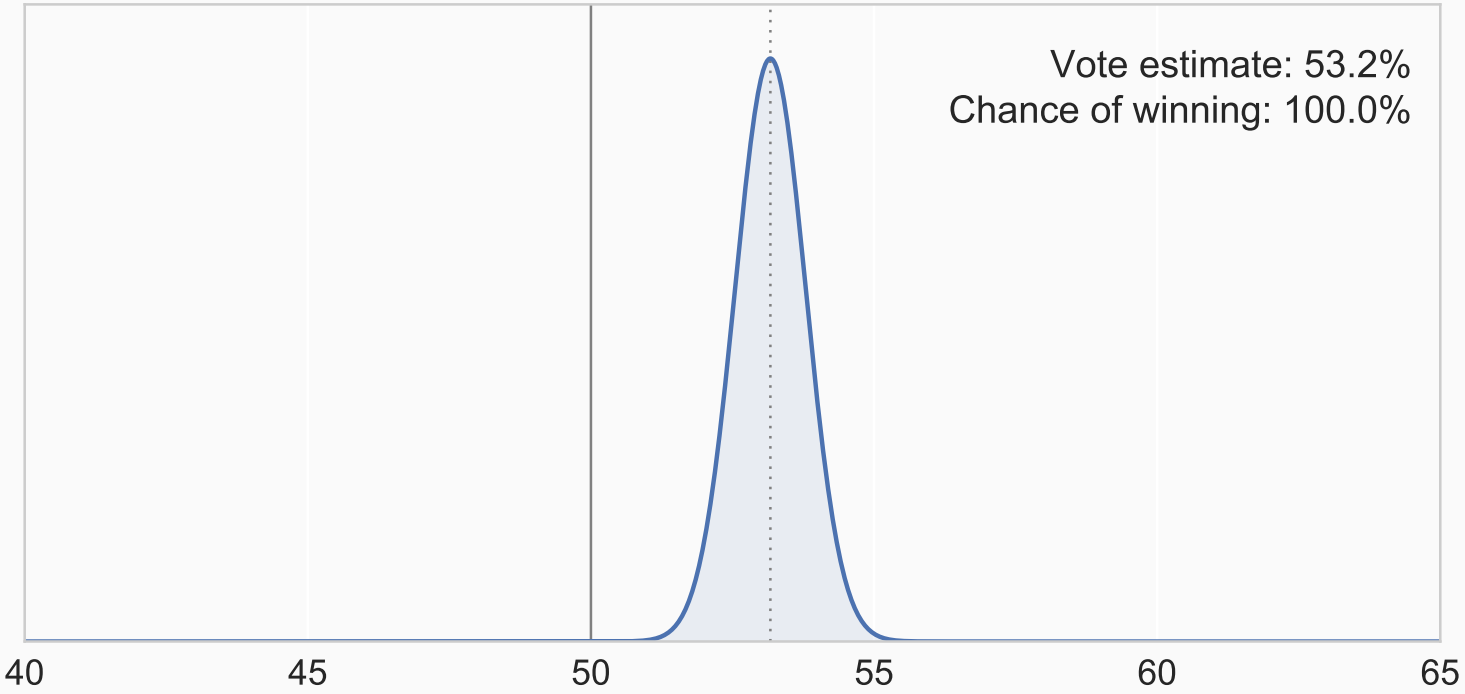
Posterior distribution, assuming simple random sampling



Date	Firm	Trump	Clinton	N
11/1/16	PPP	41%	48%	891
11/1/16	Loras College	38%	44%	500
11/2/16	Remington	41%	49%	2720
11/6/16	Ipsos/Reuters	40%	46%	842
11/6/16	UPI/CVOTER	43%	51%	372

Clinton vote share, Wisconsin

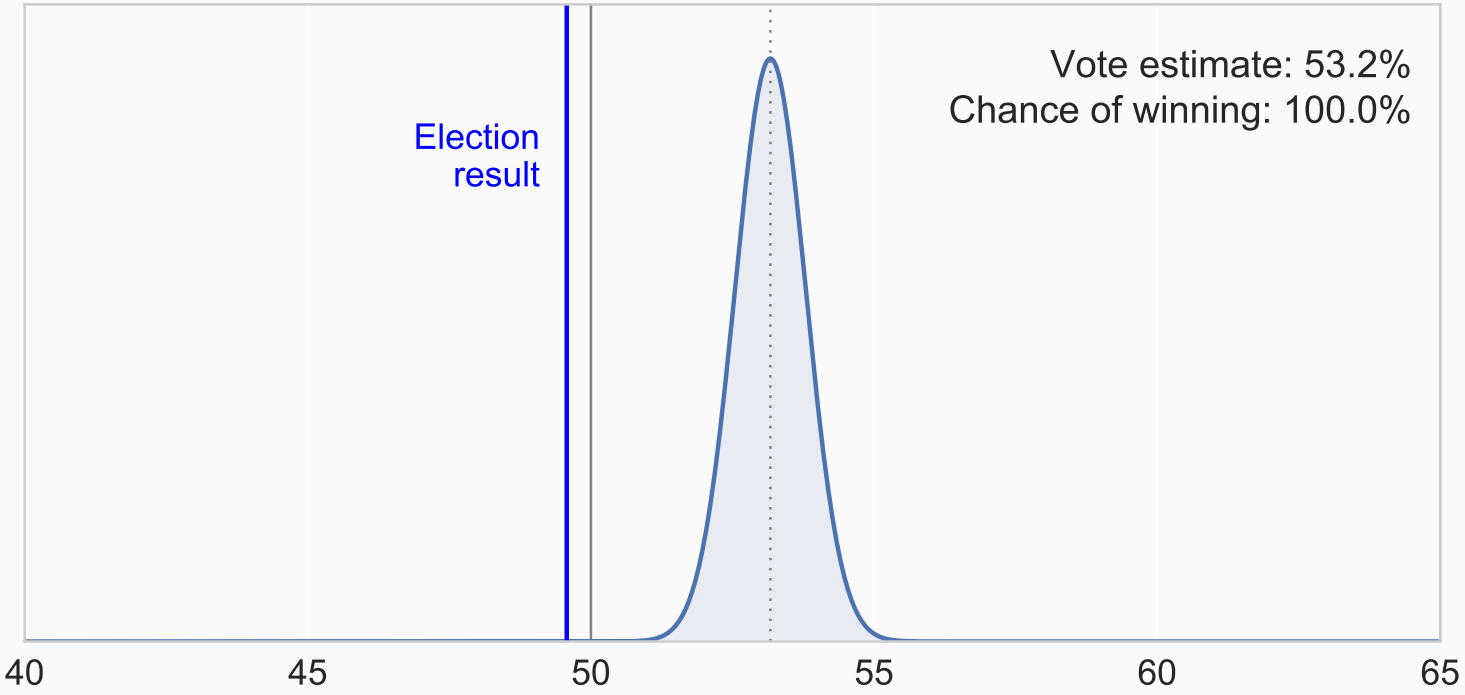
Posterior distribution, assuming simple random sampling



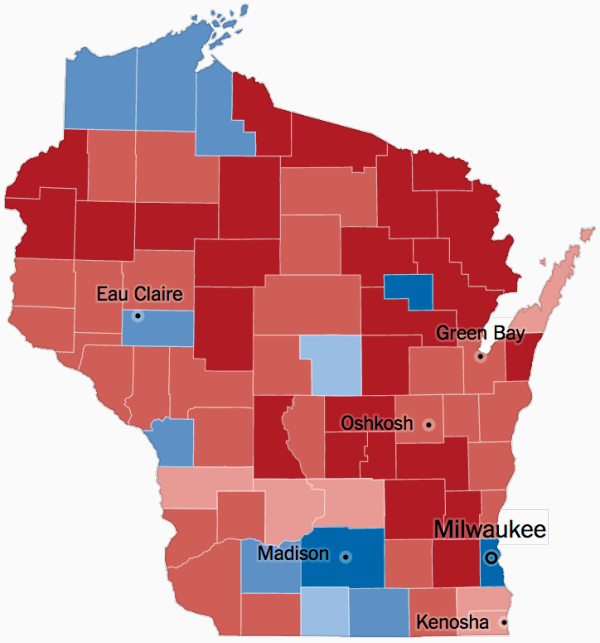
Date	Firm	Trump	Clinton	N
11/1/16	PPP	41%	48%	891
11/1/16	Loras College	38%	44%	500
11/2/16	Remington	41%	49%	2720
11/6/16	Ipsos/Reuters	40%	46%	842
11/6/16	UPI/CVOTER	43%	51%	372
11/7/16	SurveyMonkey	42%	44%	2246

Clinton vote share, Wisconsin

Posterior distribution, assuming simple random sampling



Date	Firm	Trump	Clinton	N
11/1/16	PPP	41%	48%	891
11/1/16	Loras College	38%	44%	500
11/2/16	Remington	41%	49%	2720
11/6/16	Ipsos/Reuters	40%	46%	842
11/6/16	UPI/CVOTER	43%	51%	372
11/7/16	SurveyMonkey	42%	44%	2246

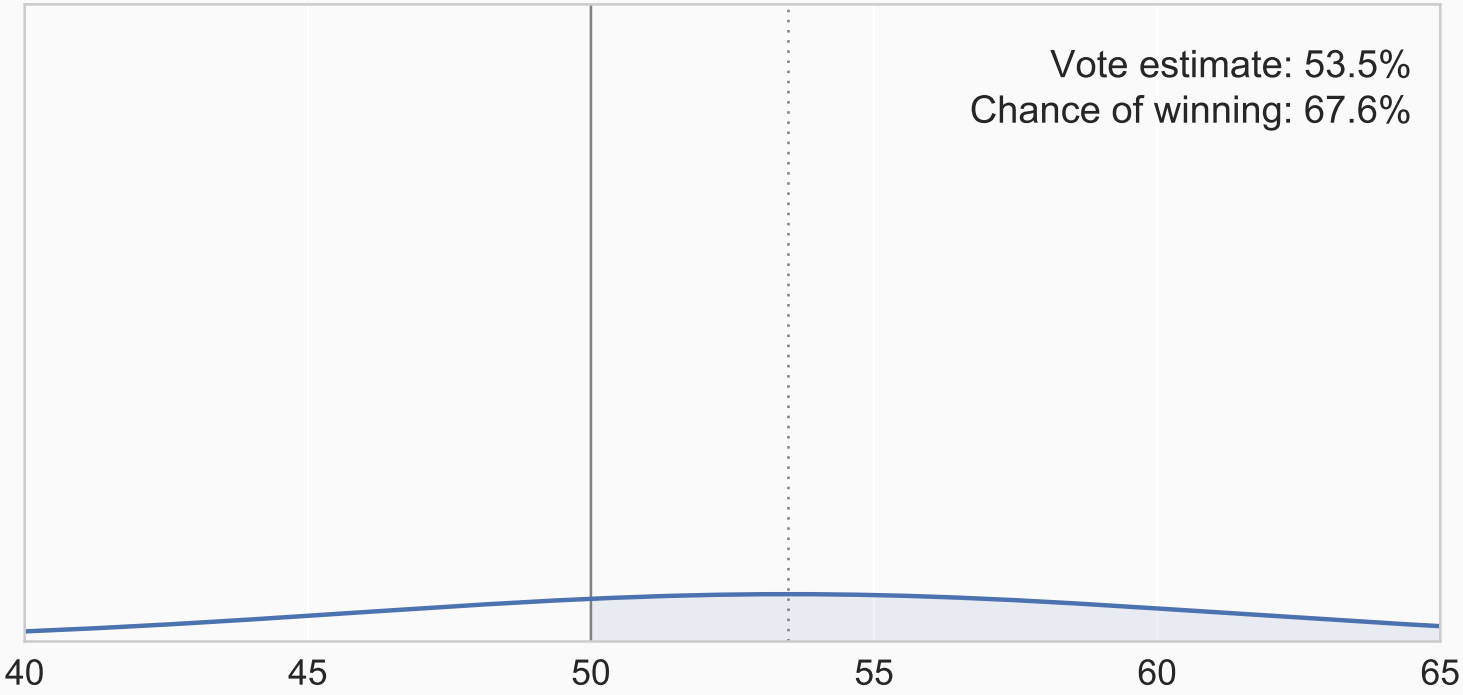


But some analysts said Clinton's chances of winning Wisconsin were as low as **93% (The Upshot)** or **84% (FiveThirtyEight)**.

- Does this mean they were more "right"?
- What assumptions would that require?

Clinton vote share, Wisconsin

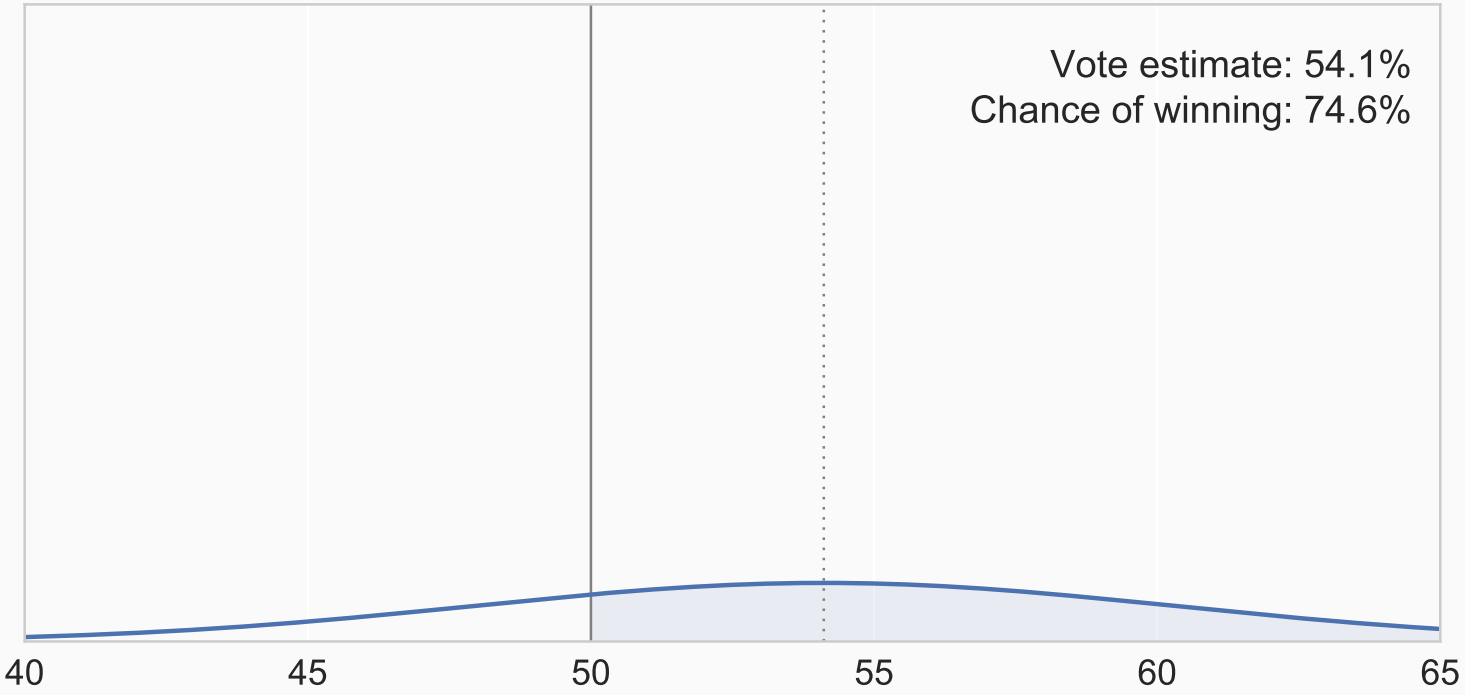
Posterior distribution, assuming one-twentieth (5%) the sample size



Date	Firm	Trump	Clinton	N	“effective” N
11/1/16	PPP	41%	48%	891	45

Clinton vote share, Wisconsin

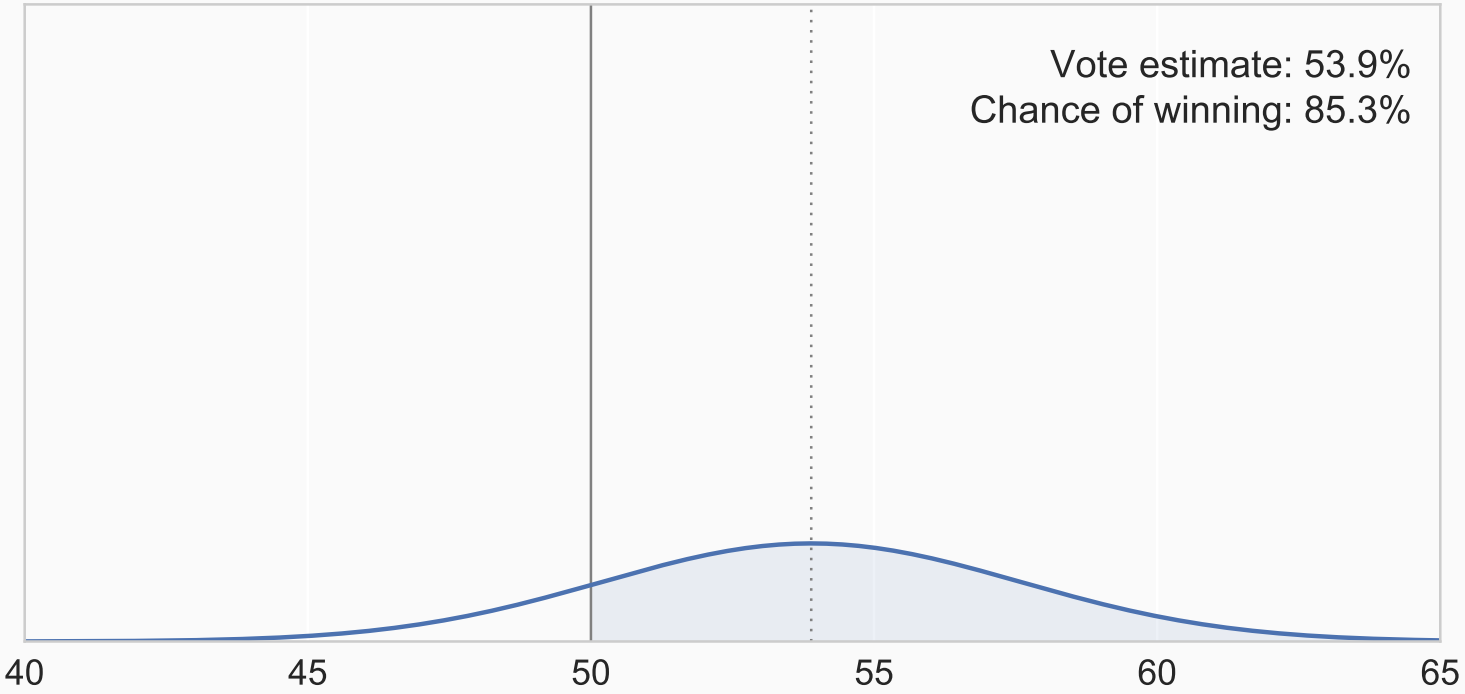
Posterior distribution, assuming one-twentieth (5%) the sample size



Date	Firm	Trump	Clinton	N	“effective” N
11/1/16	PPP	41%	48%	891	45
11/1/16	Loras College	38%	44%	500	25

Clinton vote share, Wisconsin

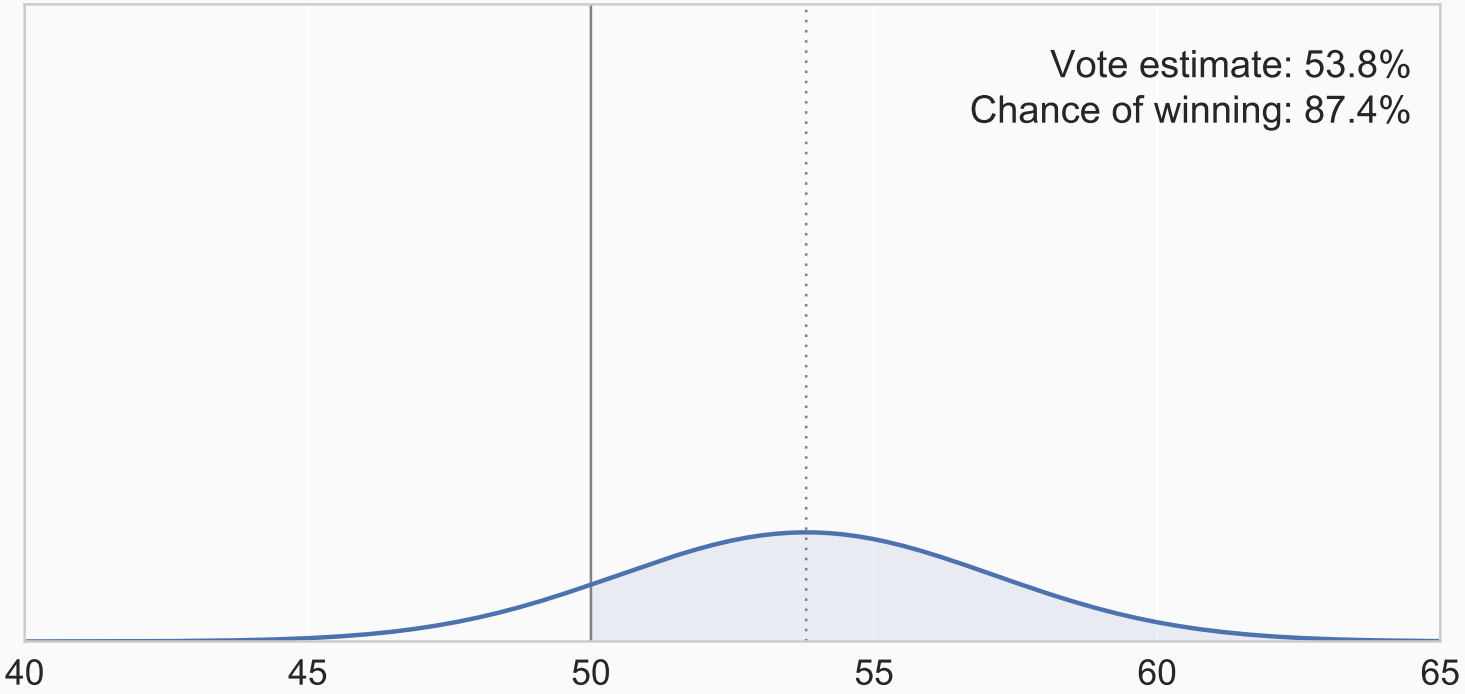
Posterior distribution, assuming one-twentieth (5%) the sample size



Date	Firm	Trump	Clinton	N	“effective” N
11/1/16	PPP	41%	48%	891	45
11/1/16	Loras College	38%	44%	500	25
11/2/16	Remington	41%	49%	2720	136

Clinton vote share, Wisconsin

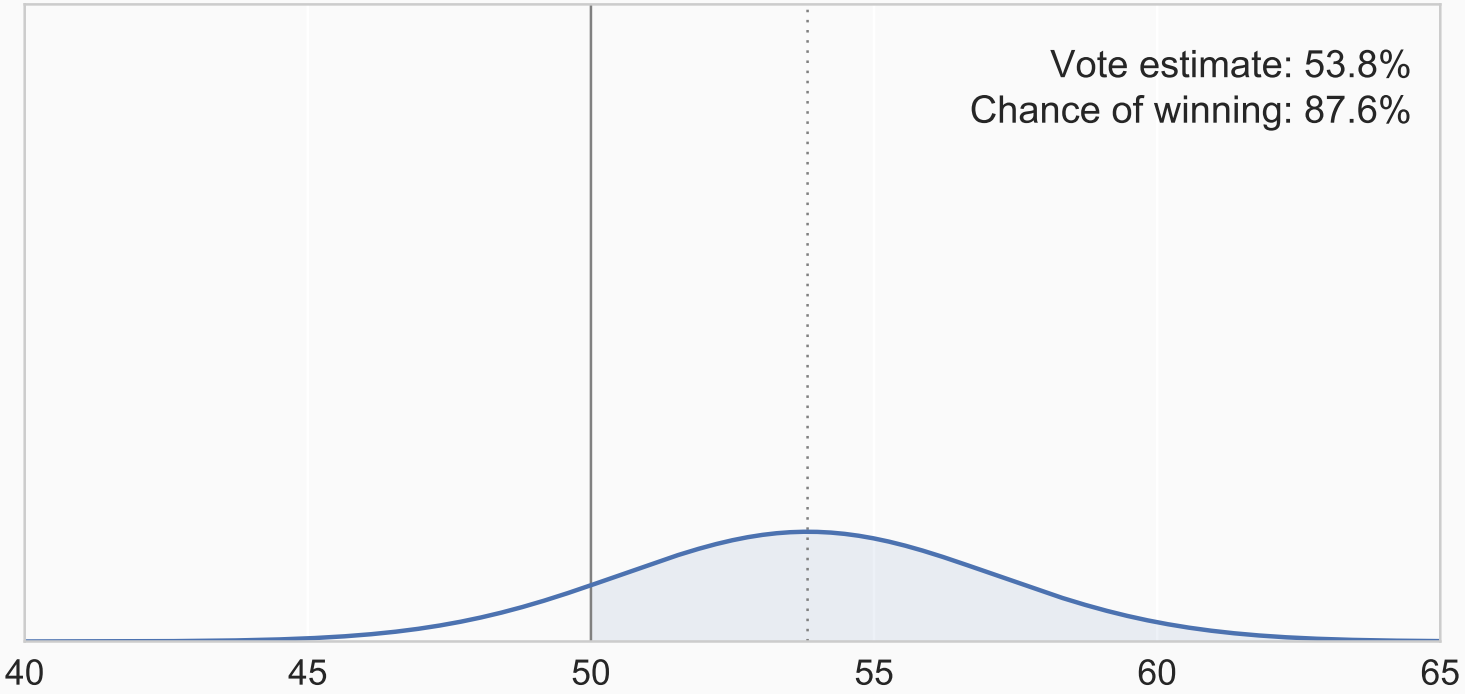
Posterior distribution, assuming one-twentieth (5%) the sample size



Date	Firm	Trump	Clinton	N	“effective” N
11/1/16	PPP	41%	48%	891	45
11/1/16	Loras College	38%	44%	500	25
11/2/16	Remington	41%	49%	2720	136
11/6/16	Ipsos/Reuters	40%	46%	842	42

Clinton vote share, Wisconsin

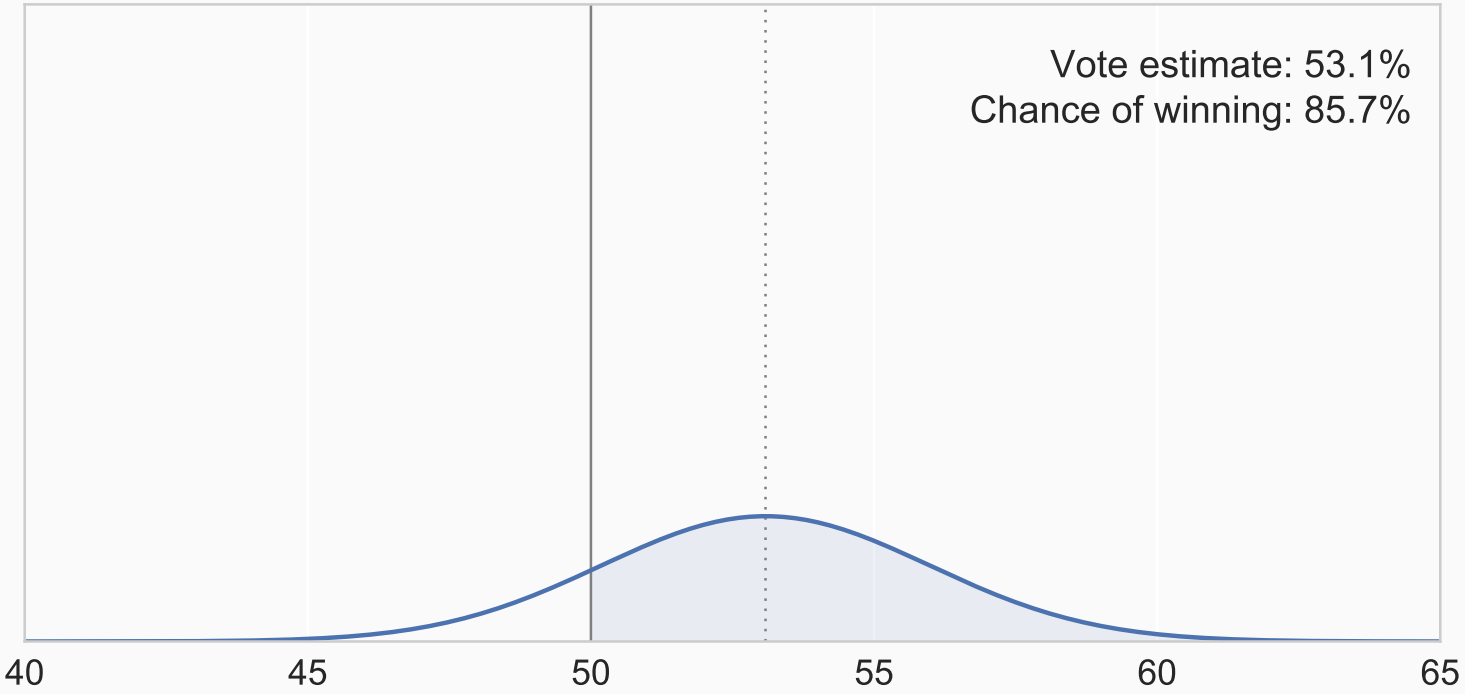
Posterior distribution, assuming one-twentieth (5%) the sample size



Date	Firm	Trump	Clinton	N	“effective” N
11/1/16	PPP	41%	48%	891	45
11/1/16	Loras College	38%	44%	500	25
11/2/16	Remington	41%	49%	2720	136
11/6/16	Ipsos/Reuters	40%	46%	842	42
11/6/16	UPI/CVOTER	43%	51%	372	19

Clinton vote share, Wisconsin

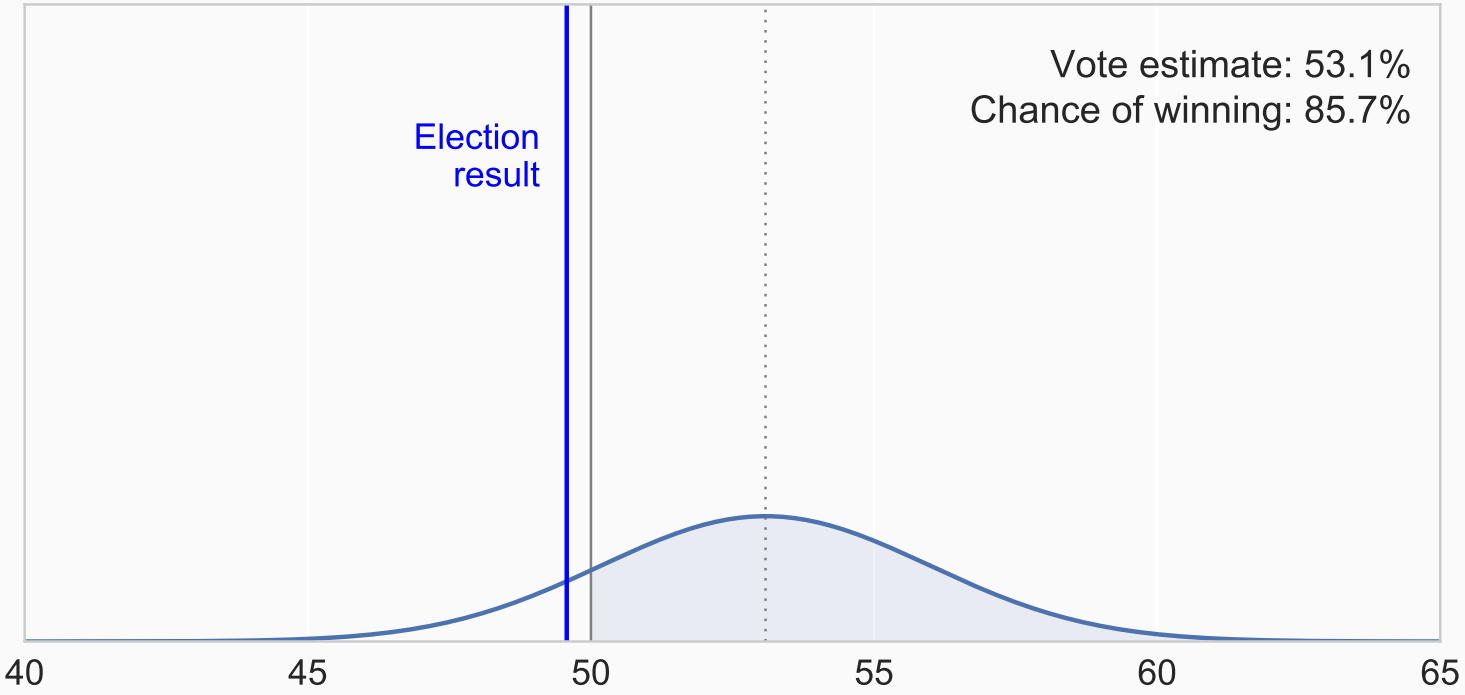
Posterior distribution, assuming one-twentieth (5%) the sample size



Date	Firm	Trump	Clinton	N	"effective" N
11/1/16	PPP	41%	48%	891	45
11/1/16	Loras College	38%	44%	500	25
11/2/16	Remington	41%	49%	2720	136
11/6/16	Ipsos/Reuters	40%	46%	842	42
11/6/16	UPI/CVOTER	43%	51%	372	19
11/7/16	SurveyMonkey	42%	44%	2246	112

Clinton vote share, Wisconsin

Posterior distribution, assuming one-twentieth (5%) the sample size



Date	Firm	Trump	Clinton	N	"effective" N
11/1/16	PPP	41%	48%	891	45
11/1/16	Loras College	38%	44%	500	25
11/2/16	Remington	41%	49%	2720	136
11/6/16	Ipsos/Reuters	40%	46%	842	42
11/6/16	UPI/CVOTER	43%	51%	372	19
11/7/16	SurveyMonkey	42%	44%	2246	112

CONCLUSIONS

- We know how to aggregate **poll means**
- Aggregating **uncertainty** is much more challenging
- The possibility of large, systematic bias suggests **heavily discounting the amount of information** in any single poll

Contact:

Drew Linzer

drew@civiqs.com
@DrewLinzer