

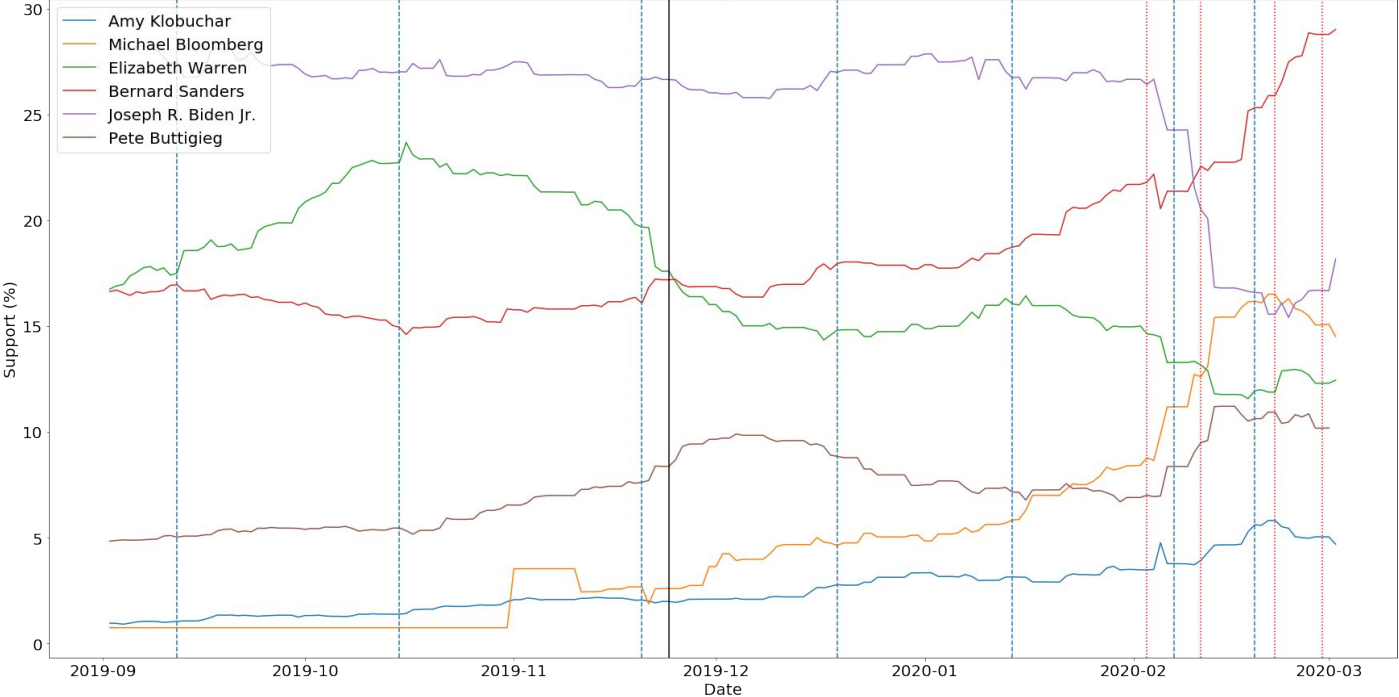


BPI Take Home

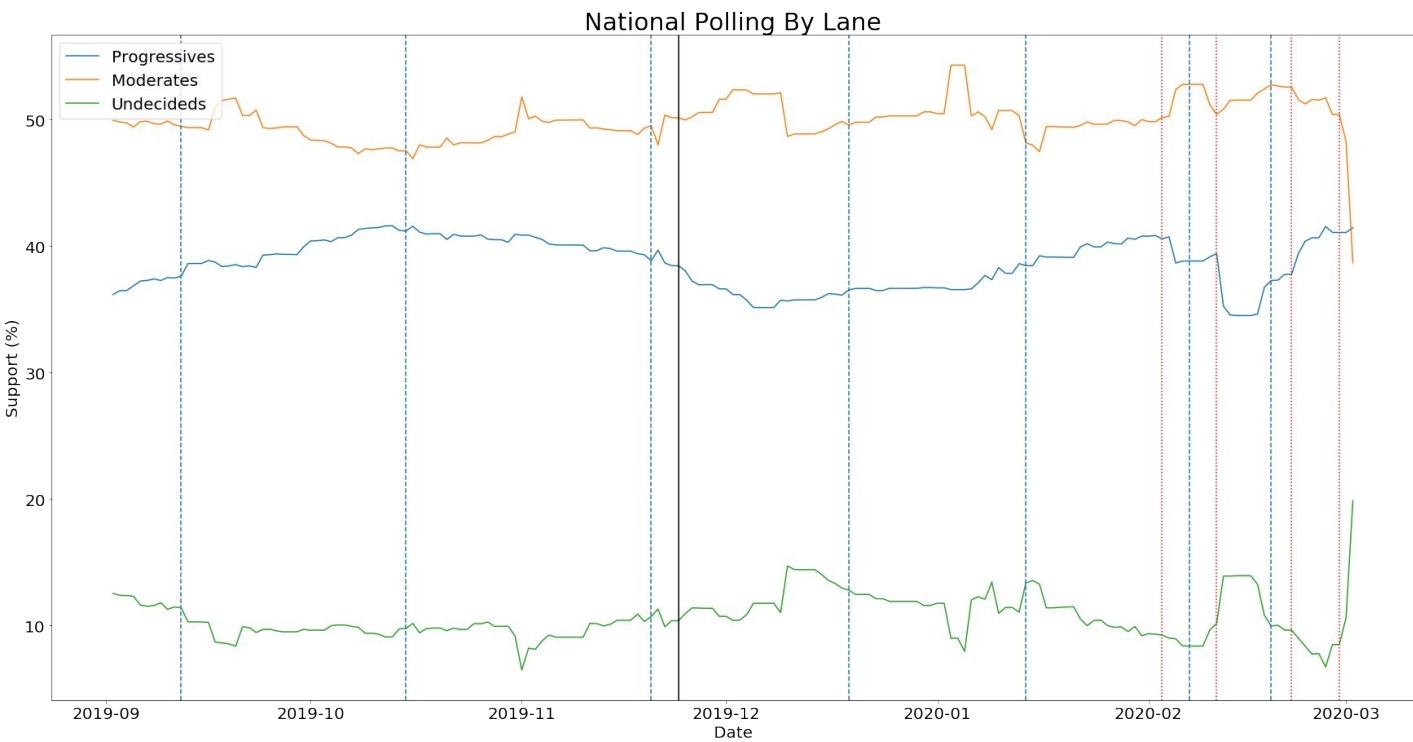
Willy Raedy



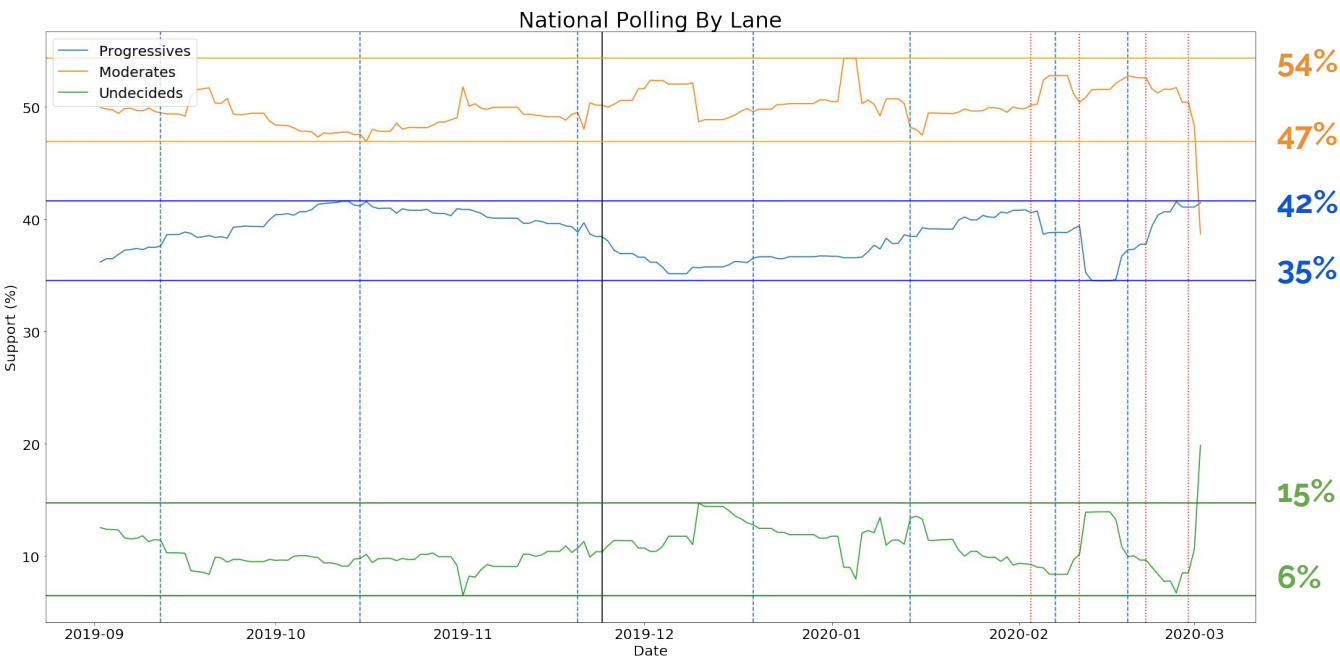
National Polling By Candidate



Fair amount of volatility



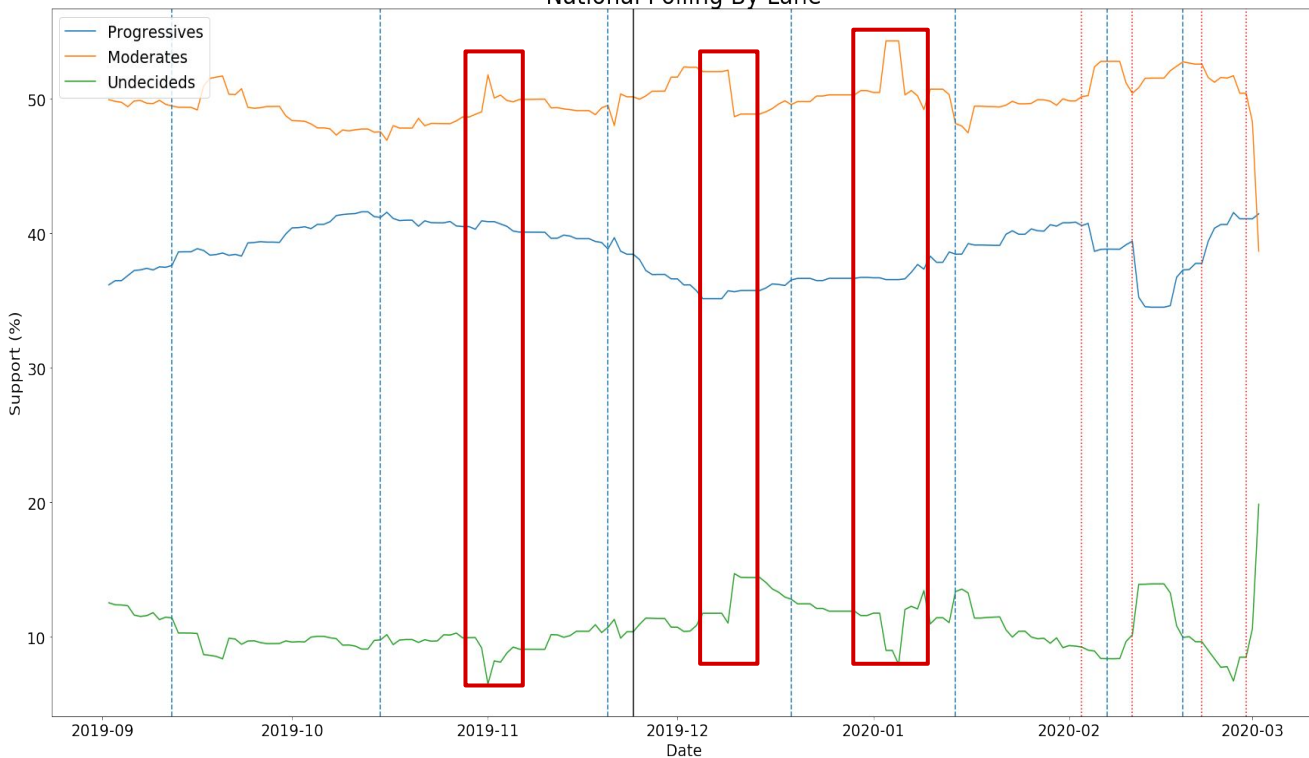
Much less volatility



Much less volatility



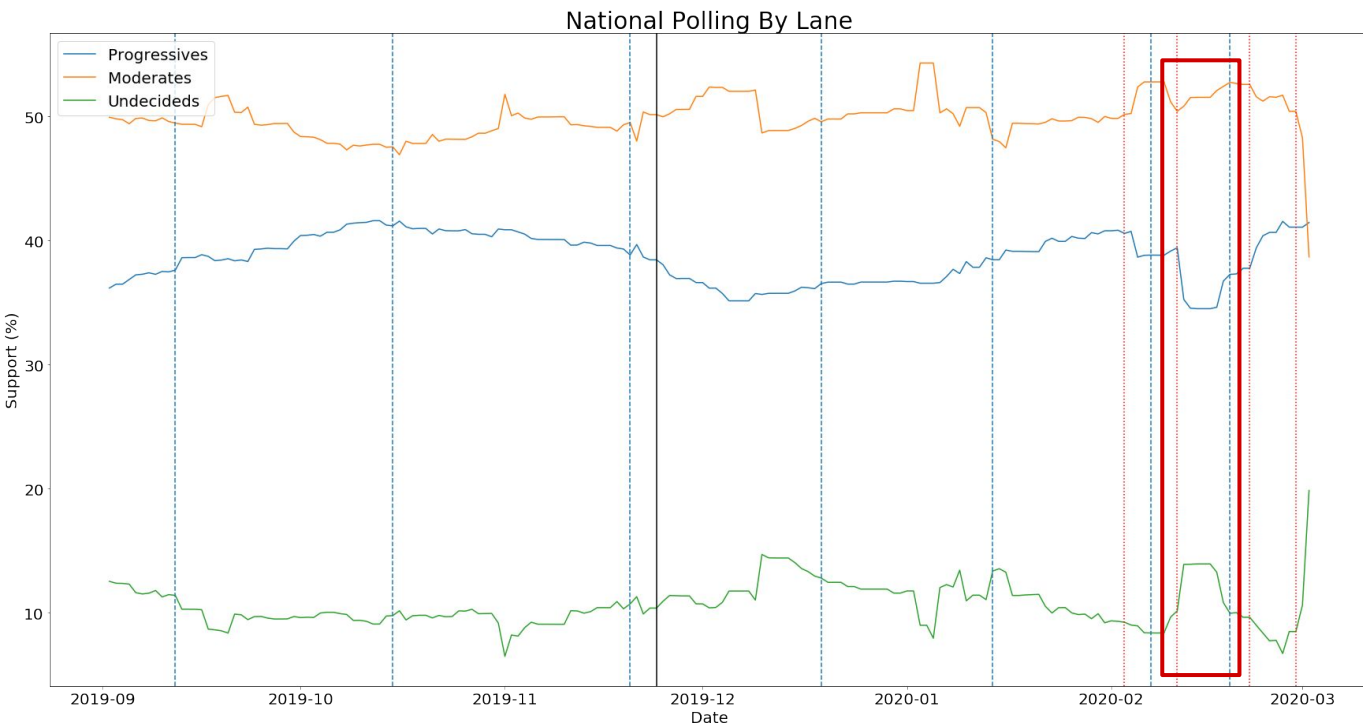
National Polling By Lane



Shocks between moderates
and undecided

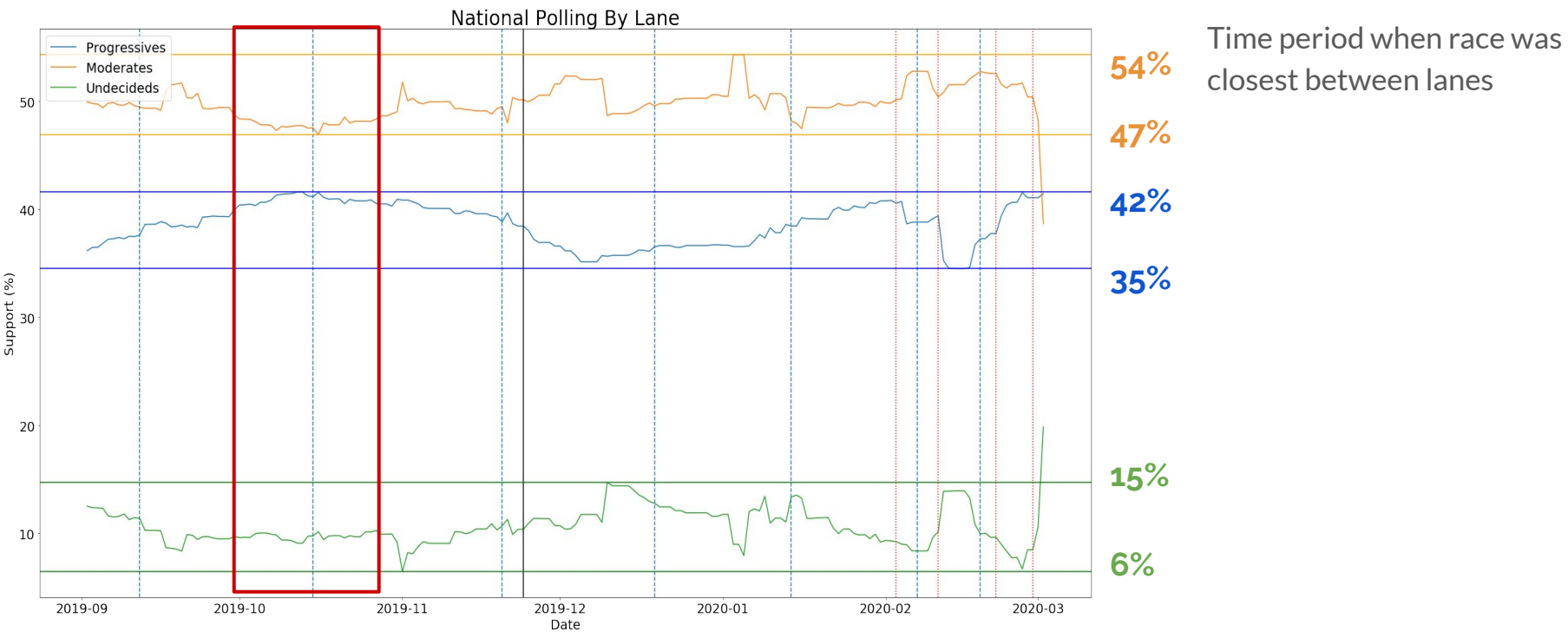
Move as distinct groups

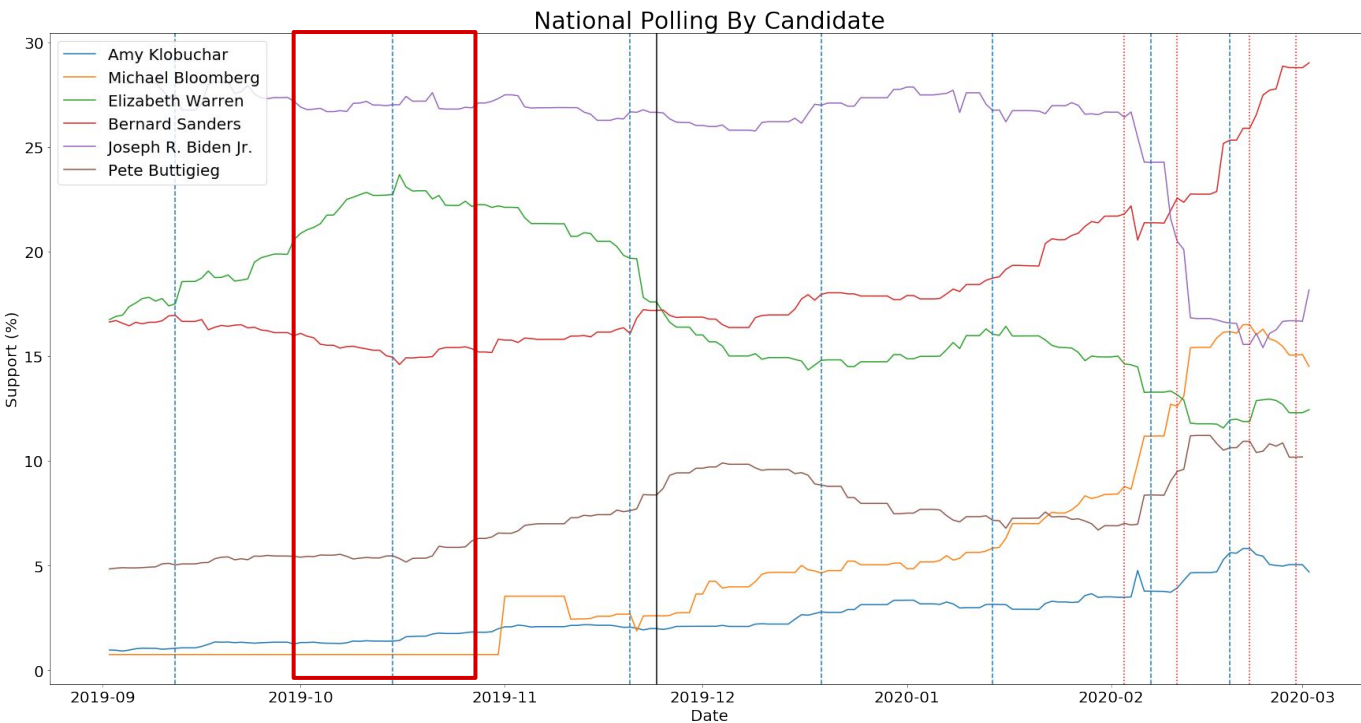
Unclear what these events are



Shocks between progressives
and undecided

Yang exits race after NH





Time period when race was closest between lanes

Corresponds to Warren's peak

Perhaps Warren was able to attract supports from moderates at this time



The “Lanes” Theory

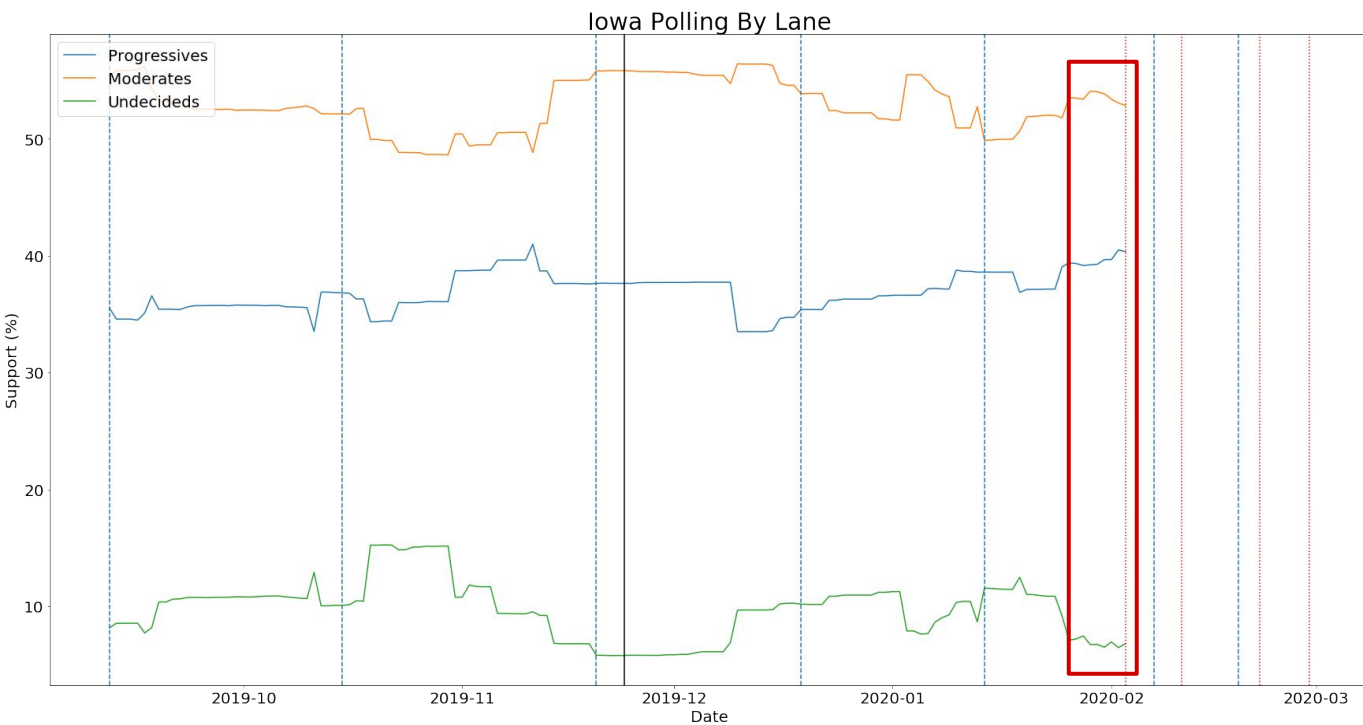
Progressive vs. Moderates split seems stable

Warren crossover appeal

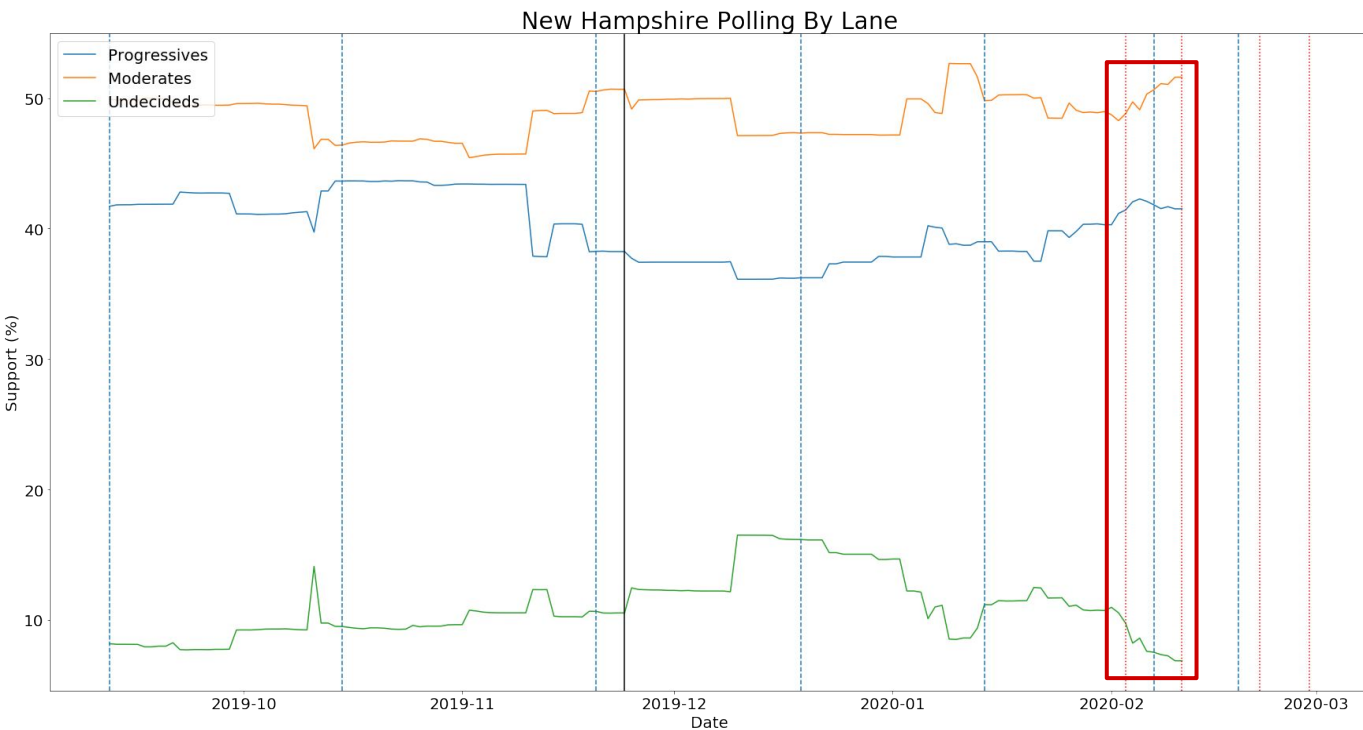
Moderates > Progressives by about 12 points

To win Progressives need:

- Moderates remain fractured
- Win undecideds at a high rate



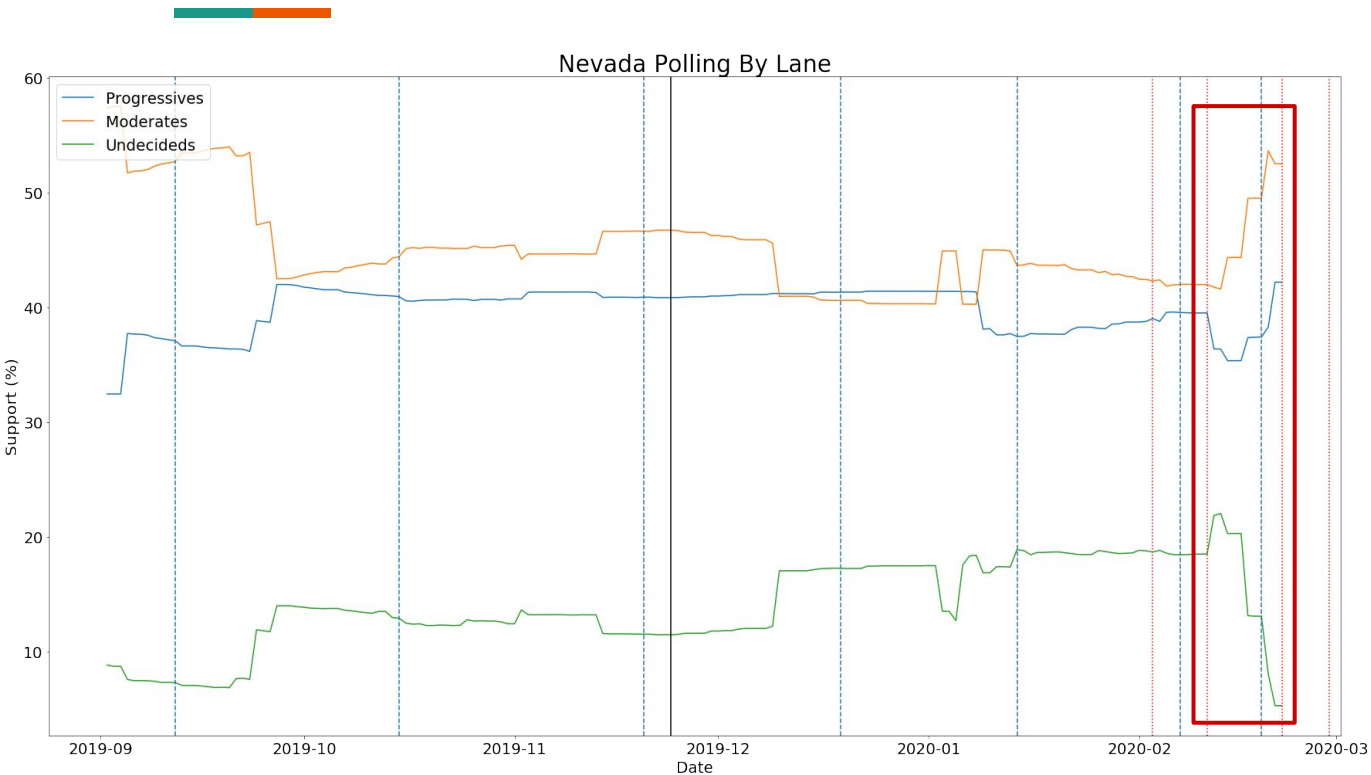
Iowa late deciders were roughly even, non-existent



Iowa late deciders were roughly even, non-existent

In three other primaries, late deciders broke for moderates

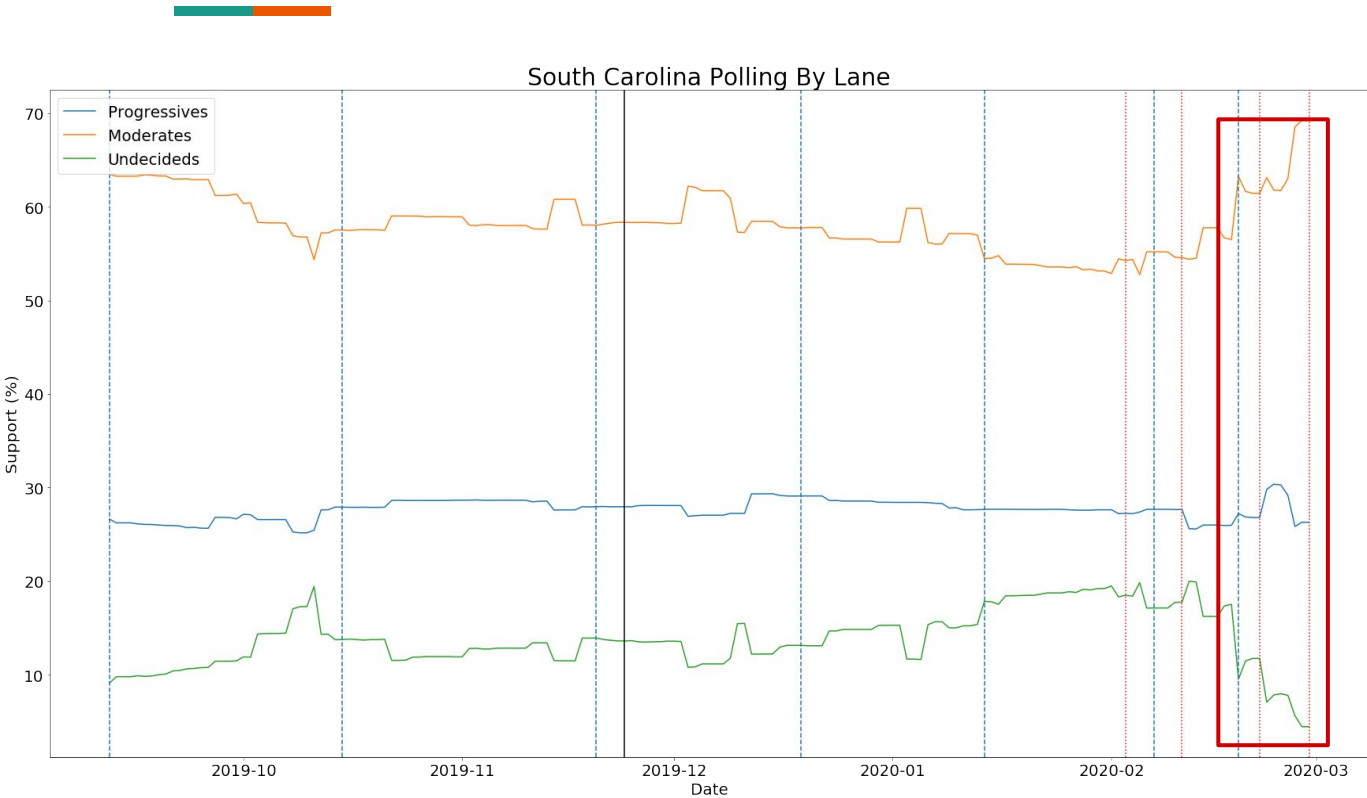
Bad sign for progressives



Iowa late deciders were roughly even, non-existent

In three other primaries, late deciders broke for moderates

Bad sign for progressives



Iowa late deciders were roughly even, non-existent

In three other primaries, late deciders broke for moderates

Bad sign for progressives

Relationship Between Media and Polls

538 found nearly 1x1 relationship

News API limited to one month, first page

- Lots of noise in last month

2474 articles from 2/3/2020 to 3/2/2020

Article labeled “about a candidate” if:

- Description included candidate’s last name
- No other candidate name

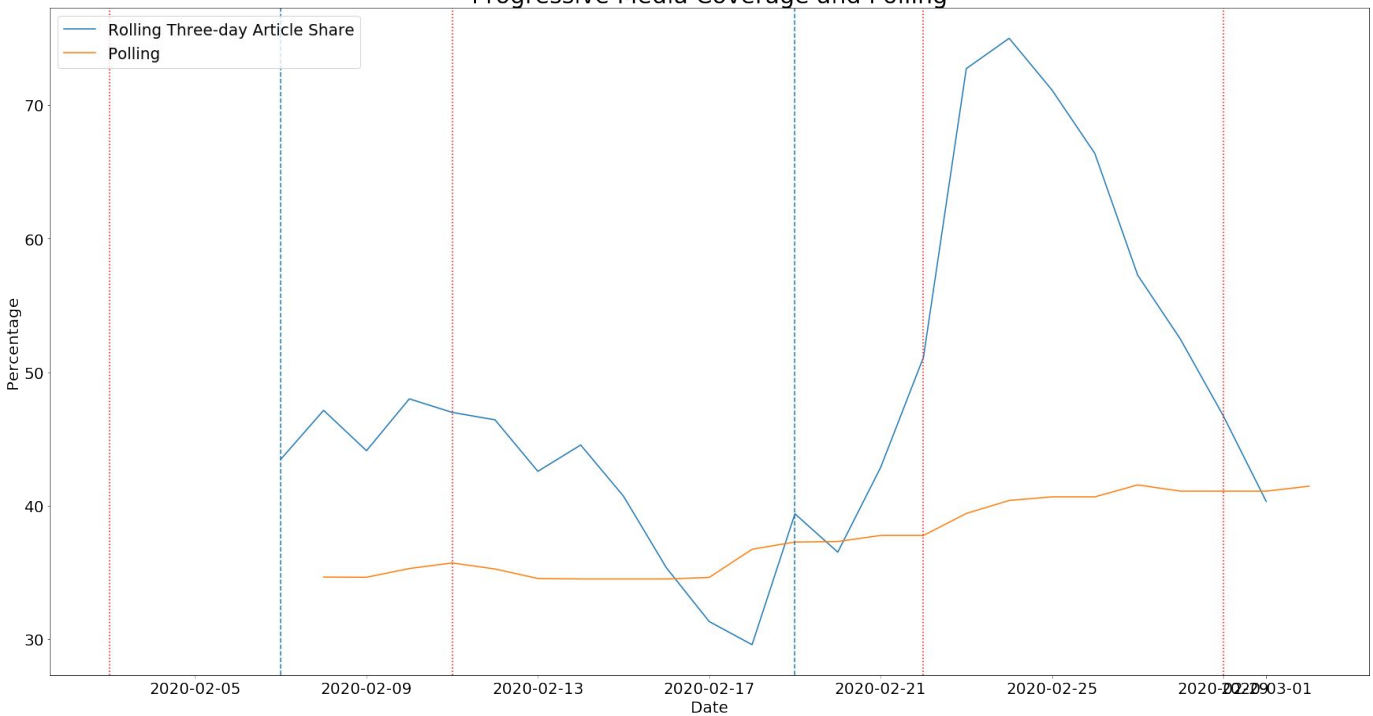
Polls can be predicted based on favorability ratings and media coverage

CANDIDATE	NET FAVORABILITY	SHARE OF MEDIA COVERAGE	ACTUAL POLL AVERAGE	PREDICTED POLL AVERAGE
Carson	+61%	7%	14%	10%
Rubio	+61%	4%	8%	9%
Cruz	+49%	5%	8%	8%
Fiorina	+42%	4%	4%	6%
Huckabee	+42%	2%	4%	6%
Bush	+29%	8%	9%	7%
Trump	+28%	54%	28%	28%
Santorum	+23%	1%	1%	3%
Kasich	+21%	5%	3%	4%
Christie	+18%	5%	3%	4%
Paul	+9%	3%	4%	2%
Gilmore	-7%	0%	0%	0%
Graham	-8%	2%	1%	0%
Pataki	-11%	0%	0%	0%

Source: [538](#)



Progressive Media Coverage and Polling



As expected, pretty noisy

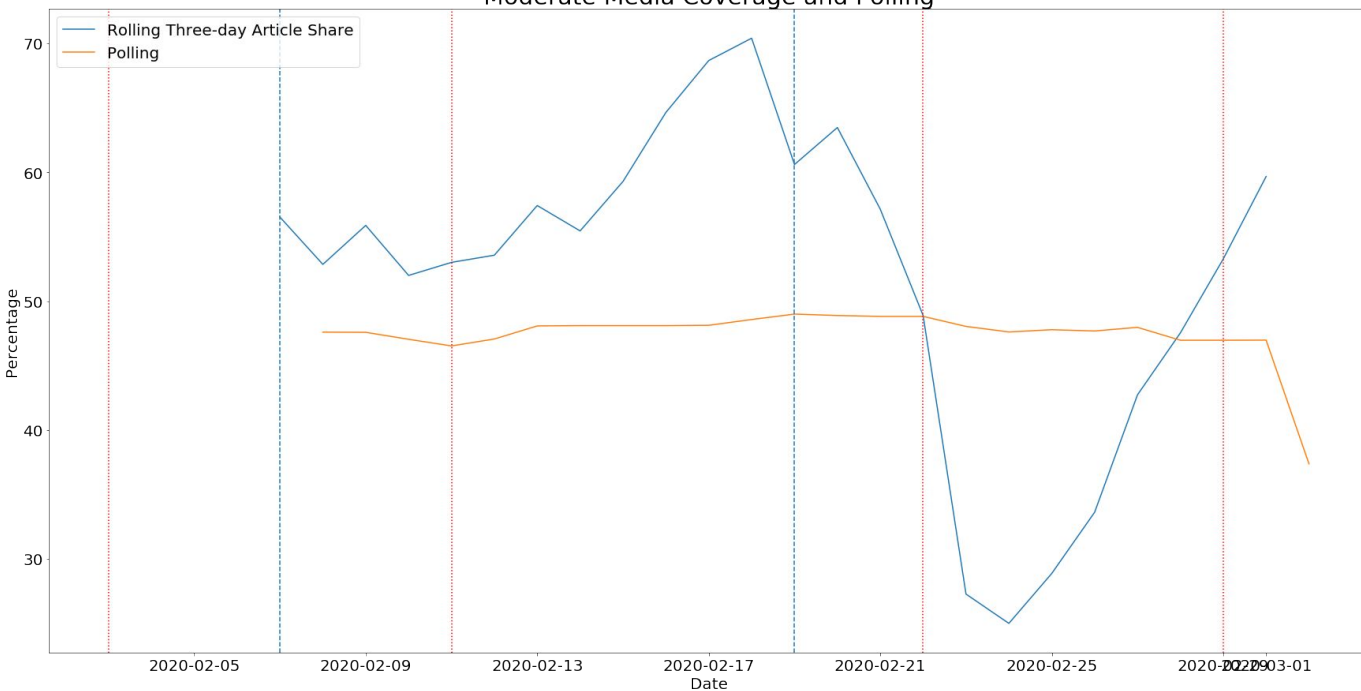
Perhaps media coverage reacts faster than polls to events

Media coverage on upswing before Sander's Nevada win

Down turn before South Carolina loss



Moderate Media Coverage and Polling



Essentially inverse of last graph

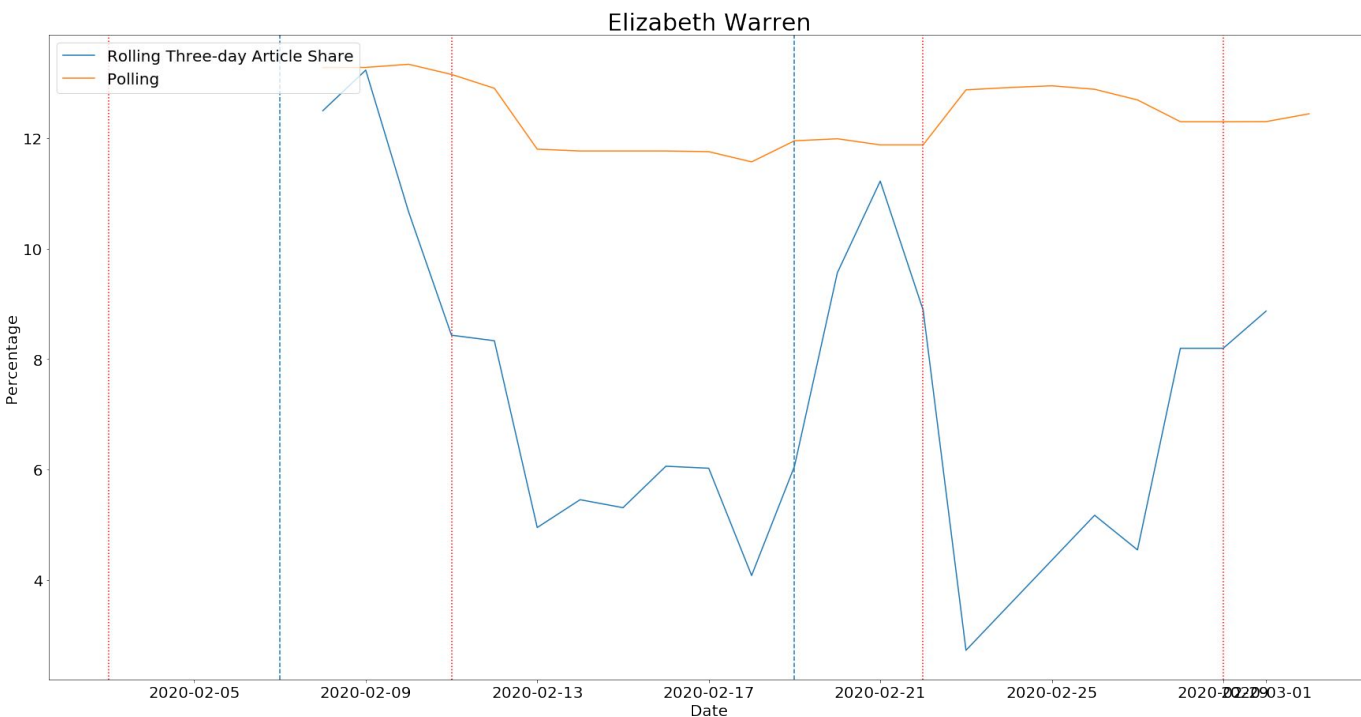


Relationship Between Media and Polls

Very small time window

Lots of noise

Changes in coverage before primaries matched performance



Warren may have legitimate gripe

Only candidate to get coverage below her polling for the entire time period

Would need more data



With more time

Collect more articles over a wider time scale

Topic modeling

- Look for associations between candidates, lanes, and topics over time

Try different sentiment and tone analyzers

- Part of speech analysis

Predict polling based on media features