Polling and Politics

Josh Clinton
Abby and Jon Winkelried Chair
Vanderbilt University
(Too much) Focus on the campaign...

- News coverage much more focused on horse-race than policy
This chart combines the latest opinion polls into trendlines using a poll-tracking model and is updated whenever a new poll is released.
Tell me again how you and your friends still have jobs?
Why are surveys important?

“Surveys are especially useful for dealing with issues of democratic representation. Participation is a mechanism for representation, a means by which the governing officials are informed of the preferences and needs of the public and are induced to respond to those preferences and needs.”

Sidney Verba
Harvard University
What is Survey Research?

Unobserved population statistic → sample → inference → Sample statistic
Early Survey Research

- Magazine from the early 1900s that conducted presidential election polls
- Surveyed subscribers
  - Millions responded to each poll
- Accurately predicted the 1920, 1924, 1928, & 1932 Presidential Elections
1936 Presidential Election

Alfred Langdon
Governor of Kansas
Republican Presidential Candidate

Franklin Roosevelt
President of the United States
Democrat Presidential Candidate
1936 Presidential Election

**Literary Digest Election Forecast**

- *Literary Digest* surveyed over 2 millions Americans.
- Forecasted Langdon would win with 57.1% of the popular vote and 370 to 161 of the Electoral College Vote.
1936 Presidential Election

Actual 1936 Election Outcome

Literary Digest: Langdon 57% Roosevelt 43%
Actual Election Result: Langdon 39% Roosevelt 61%
Who got the 1936 Election right?

• Was able to accurately forecast the outcome of the 1936 election within 1% of the actual result

• Instead of over 2 million voters, Gallup interviewed approximately 50,000

• Why? Survey Sampling
Main findings

Did polls fail in 2016?

• There are two storylines to the performance of polls in the 2016 general election
  – National polls did well, historically quite accurate
  – State-level polls had historically bad year; Some large errors in the Upper Midwest
National General Election Polls’ Performance

Polls included in the analysis:

- Final field dates falling within 13 days of Election Day (Oct. 26th or later) if field period began by Oct. 16th.
- Only the final poll from each firm
- 23 final national polls of the 39 total during that period
- 23 polling organizations/sponsor combinations
Measuring Error

Absolute Error

• Absolute value of
  ((the margin (%Clinton-%Trump) in the poll)
  minus
  (the margin (%Clinton-%Trump) in certified vote))

• Poll: Clinton 46% - Trump 42%. 4 percentage points
• Vote: Clinton 48.2% - Trump 46.1%. 2.1 percentage points
• ABS(4 – 2.1) = 1.9
How did the national polls do?

• Average absolute error of 2.2 percentage points
• Average signed error of 1.3 percentage points
• Among the best performances since 1936
National Poll Errors Over The Years

Orange line: Average error in vote margin
Bars: Average bias in vote margin (Blue/Positive overestimates Democratic vote margin, Red/Negative overestimates Republicans)

2016 data based on national polls that were completed within 13 days of Election Day. Previous years based on National Council for Public Polls analysis of final poll estimates, some occurring before the 13-day period.
STATE General Election Polls’ Performance

Polls included in the analysis:

• Final field dates falling within 13 days of Election Day (Oct. 26th or later) if field period began by Oct. 16th
• ALL polls from each firm within that period
• 423 final state polls
• 81 polling organizations/sponsor combinations
• Same error calculations as national polls
How did the state polls do?

- Average absolute error of 5.1 percentage points
- Average signed error of 3.0 percentage points
- The poorest performance at the state level compared to the last four elections (2000-2012)
State Poll Errors Over The Years

Orange line represents average absolute error
Bars represent average signed error
Red bars indicate overestimation of Republican vote margin
Blue bars indicate overestimation of Democratic vote margin
Main findings
What caused under-estimation of support for Trump?

• Evidence of late change in vote preferences
  – Two historically unpopular candidates. Many voter liked neither.
  – Late change concentrated in key states, not national

<table>
<thead>
<tr>
<th></th>
<th>% Voters who decided in final week</th>
<th>Vote choice among voters deciding in final week</th>
<th>Vote choice among voters deciding earlier</th>
<th>Estimated Trump gain from late deciders</th>
<th>Election (%Trump-%Clinton)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>11%</td>
<td>55% 38%</td>
<td>48% 49%</td>
<td>2.0%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Michigan</td>
<td>13%</td>
<td>50% 39%</td>
<td>48% 48%</td>
<td>1.4%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>15%</td>
<td>54% 37%</td>
<td>50% 48%</td>
<td>2.3%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>14%</td>
<td>59% 30%</td>
<td>47% 48%</td>
<td>4.3%</td>
<td>0.8%</td>
</tr>
<tr>
<td>National</td>
<td>13%</td>
<td>45% 42%</td>
<td>46% 49%</td>
<td>0.8%</td>
<td>-2.1%</td>
</tr>
</tbody>
</table>

Note – Analysis from Aaron Blake (2016) using NEP exit poll data.
Because of Comey?

Hillary Clinton's national polling lead in the months before the election if the race was judged based on …

... when polls were taken

... when polls were released

0 pts. 2 pts. 4 pts. +6 pts.

Mon 15 Sep. Thu 15 Oct. Sat 15 Fri 28 Tue 08

Comey letter
Sources of Error in Surveys

• When building instruments and administering surveys, there are many potential sources of error and bias
  • Administration Based:
    – Sampling Bias
    – Interviewer Effects
  
  • Respondent and Societal Based:
    – Social Desirability
    – Contextual Effects
  
  • Instrument Based:
    – Question Order Effects
    – Question Wording
    – Response Options
Survey Sampling

- **Survey Sampling**: Process of selecting elements from a target population (e.g. voters) to conduct a survey.
Contemporary Sampling Bias: Land Lines & Cell Phones

- Why might these trends be a problem for phone surveys?
Main findings

What caused under-estimation of support for Trump?

• Most state-level polls did not adjust for overrepresentation of college graduates, which was correlated with Clinton support
Main findings
What caused under-estimation of support for Trump?

- Most state-level polls did not adjust for overrepresentation of college graduates, which was correlated with Clinton support.

Source: NEP national Exit Poll 2012, 2016
Main findings

What caused under-estimation of support for Trump?

- Most state-level polls did not adjust for overrepresentation of college graduates, which was correlated with Clinton support.

<table>
<thead>
<tr>
<th>Share of Pollsters That Adjusted on Education in Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Poll</td>
</tr>
<tr>
<td>Michigan polls</td>
</tr>
<tr>
<td>Wisconsin polls</td>
</tr>
<tr>
<td>North Carolina polls</td>
</tr>
<tr>
<td>Florida polls</td>
</tr>
<tr>
<td>Pennsylvania polls</td>
</tr>
<tr>
<td>Ohio polls</td>
</tr>
<tr>
<td>National polls</td>
</tr>
</tbody>
</table>

Note: Weighting information was missing for 23 polls, which were all imputed as not weighting on education, based on information among similar polls that did disclose their weighting variables.
Effects of Weighting on Education

Figure 10. Poll Estimates with and Without Weighting Adjustment for Education, Relative to 2016 Presidential Vote Outcome. Source – University of New Hampshire poll conducted November 3-6, 2016 with 707 likely voters. Michigan State University poll conducted September 1 – October 30, 2016 with 743 likely voters.
Who Votes?
Change in turnout (2012 to 2016)
by county nationwide by vote share for Barack Obama in 2012

Source – County-level vote data come from uselectionatlas.org
## County-level change in turnout in Michigan

<table>
<thead>
<tr>
<th>Counties Where Obama's 2012 Vote Percentage Was...</th>
<th>Percent Change</th>
<th>2012 Composition</th>
<th>2016 Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;35%</td>
<td>8.5%</td>
<td>2.9%</td>
<td>3.1%</td>
</tr>
<tr>
<td>35-40%</td>
<td>6.4%</td>
<td>5.2%</td>
<td>5.5%</td>
</tr>
<tr>
<td>40-45%</td>
<td>3.9%</td>
<td>10.2%</td>
<td>10.4%</td>
</tr>
<tr>
<td>45-50%</td>
<td>3.4%</td>
<td>19.1%</td>
<td>19.3%</td>
</tr>
<tr>
<td>50-54%</td>
<td>2.8%</td>
<td>27.6%</td>
<td>27.8%</td>
</tr>
<tr>
<td>55-59%</td>
<td>1.0%</td>
<td>7.0%</td>
<td>6.9%</td>
</tr>
<tr>
<td>60-65%</td>
<td>-0.3%</td>
<td>7.0%</td>
<td>6.8%</td>
</tr>
<tr>
<td>65%+</td>
<td>-2.3%</td>
<td>21.0%</td>
<td>20.2%</td>
</tr>
<tr>
<td>100.0%</td>
<td></td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Source – County-level vote data come from uselectionatlas.org
How Clinton’s margin in swing states would have differed under 2012 turnout rates (Fraga, McElwee, Rhodes & Schaffner, 2017)

Estimates created using turnout data from Catalist and vote share data from the Cooperative Congressional Election Study. Figure by Bernard Fraga, Sean McElwee, Jesse Rhodes, and Brian Schaffner.
107k
Sources of Error in Surveys

• When building instruments and administering surveys, there are many potential sources of error and bias
  • Administration Based:
    – Sampling Bias
    – Interviewer Effects
  • Respondent and Societal Based:
    – Social Desirability
    – Contextual Effects
  • Instrument Based:
    – Question Order Effects
    – Question Wording
    – Response Options
Interviewer Effects

• **Interviewer Effect**: Any effect on data gathered from interviewing people that is caused by the behavior or characteristics of the interviewer.

<table>
<thead>
<tr>
<th>Proposed national apology for slavery</th>
<th>Black Interviewer</th>
<th>Non-black interviewer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favor</td>
<td>41%</td>
<td>33%</td>
</tr>
<tr>
<td>Oppose</td>
<td>51%</td>
<td>59%</td>
</tr>
<tr>
<td>Don’t know</td>
<td>8%</td>
<td>8%</td>
</tr>
</tbody>
</table>

*Source: May 1998 Pew Research Center Survey*
Context Effects: Godfather’s Pizza

Herman Cain
2012 Republican Presidential Candidate
Godfather’s Pizza Index by Political Party

- Democrat
- Republican
- Independent
Who Won World War II?

Sondage en France : "Quelle est, selon vous, la nation qui a le plus contribué à la défaite de l'Allemagne en 1945 ?" (Source : sondages IFOP 1945, 1994, 2004)

- Mai 1945: 57% (URSS), 20% (États-Unis), 12% (Grande-Bretagne)
- Mai 1994: 49% (États-Unis), 25% (URSS), 16% (Grande-Bretagne)
- Juin 2004: 58% (États-Unis), 20% (URSS), 16% (Grande-Bretagne)

Another example from the 2012 Debates...

ROMNEY: What things would I cut from spending? Well, first of all, I will eliminate all programs by this test, if they don't pass it: Is the program so critical it's worth borrowing money from China to pay for it? And if not, I'll get rid of it. Obamacare's on my list.

I apologize, Mr. President. I use that term with all respect, by the way.

OBAMA: I like it.

ROMNEY: Good. OK, good. So I'll get rid of that.
Opinion of “Affordable Care Act” or “Obamacare”?
Among TN Registered Voters

Source: May 2013 Vanderbilt Poll
Moving forward

• The factors that best explain polling errors in 2016 are fixable...
  – Weight to correct education imbalance (if not already doing so)
  – Field closer to election day and/or remind people of the possibility of late change
  – Do not necessarily assume the prior election is the best turnout model for current one

• ...but, there is little to stop this from happening again
  – Many state polls are under-budgeted, and that will likely continue
  – Many of these questions are unknowable – who is going to vote in 2018?
Florida Democratic Primary

<table>
<thead>
<tr>
<th>Poll</th>
<th>Date</th>
<th>Sample</th>
<th>MoE</th>
<th>Graham</th>
<th>Levine</th>
<th>Greene</th>
<th>Gillum</th>
<th>King</th>
<th>Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final Results</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>31.4</td>
<td>20.3</td>
<td>10.1</td>
<td>34.3</td>
<td>2.5</td>
<td>Gillum +2.9</td>
</tr>
<tr>
<td>RCP Average</td>
<td>7/23 - 8/25</td>
<td>--</td>
<td>--</td>
<td>26.0</td>
<td>18.8</td>
<td>14.5</td>
<td>12.0</td>
<td>6.3</td>
<td>Graham +7.2</td>
</tr>
<tr>
<td>Gravis</td>
<td>8/21 - 8/25</td>
<td>531 LV</td>
<td>4.3</td>
<td>26</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>5</td>
<td>Graham +7</td>
</tr>
<tr>
<td>Gravis</td>
<td>8/21 - 8/22</td>
<td>308 LV</td>
<td>5.6</td>
<td>26</td>
<td>18</td>
<td>19</td>
<td>15</td>
<td>5</td>
<td>Graham +7</td>
</tr>
<tr>
<td>Florida Atlantic University</td>
<td>8/16 - 8/20</td>
<td>239 LV</td>
<td>6.3</td>
<td>29</td>
<td>17</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>Graham +12</td>
</tr>
<tr>
<td>SurveyUSA</td>
<td>8/10 - 8/13</td>
<td>631 LV</td>
<td>5.2</td>
<td>22</td>
<td>22</td>
<td>16</td>
<td>11</td>
<td>3</td>
<td>Tie</td>
</tr>
<tr>
<td>Mason-Dixon</td>
<td>7/23 - 7/25</td>
<td>500 LV</td>
<td>4.5</td>
<td>27</td>
<td>18</td>
<td>12</td>
<td>10</td>
<td>7</td>
<td>Graham +9</td>
</tr>
<tr>
<td>Florida Atlantic University</td>
<td>7/20 - 7/21</td>
<td>271 RV</td>
<td>5.9</td>
<td>20</td>
<td>16</td>
<td>14</td>
<td>7</td>
<td>9</td>
<td>Graham +4</td>
</tr>
<tr>
<td>Gravis</td>
<td>7/13 - 7/14</td>
<td>1540 LV</td>
<td>2.2</td>
<td>27</td>
<td>17</td>
<td>18</td>
<td>10</td>
<td>–</td>
<td>Graham +9</td>
</tr>
<tr>
<td>NBC News/Marist</td>
<td>6/17 - 6/21</td>
<td>344 RV</td>
<td>6.5</td>
<td>17</td>
<td>19</td>
<td>4</td>
<td>8</td>
<td>3</td>
<td>Levine +2</td>
</tr>
<tr>
<td>Florida Atlantic University</td>
<td>5/4 - 5/7</td>
<td>372 RV</td>
<td>5.0</td>
<td>15</td>
<td>16</td>
<td>–</td>
<td>6</td>
<td>10</td>
<td>Levine +1</td>
</tr>
<tr>
<td>Gravis</td>
<td>2/26 - 3/19</td>
<td>775 LV</td>
<td>3.5</td>
<td>9</td>
<td>13</td>
<td>2</td>
<td>11</td>
<td>2</td>
<td>Levine +2</td>
</tr>
<tr>
<td>Gravis</td>
<td>2/1 - 2/18</td>
<td>687 RV</td>
<td>3.7</td>
<td>12</td>
<td>12</td>
<td>1</td>
<td>9</td>
<td>2</td>
<td>Tie</td>
</tr>
<tr>
<td>Mason-Dixon</td>
<td>1/29 - 2/1</td>
<td>500 LV</td>
<td>4.5</td>
<td>20</td>
<td>17</td>
<td>–</td>
<td>10</td>
<td>4</td>
<td>Graham +3</td>
</tr>
</tbody>
</table>
Closer to Home...

<table>
<thead>
<tr>
<th>Poll</th>
<th>Date</th>
<th>Sample</th>
<th>MoE</th>
<th>Blackburn (R)</th>
<th>Bredesen (D)</th>
<th>Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCP Average</td>
<td>9/11 - 10/11</td>
<td>--</td>
<td>--</td>
<td>49.3</td>
<td>43.8</td>
<td>Blackburn +5.5</td>
</tr>
<tr>
<td>NY Times/Siena</td>
<td>10/8 - 10/11</td>
<td>593 LV</td>
<td>4.2</td>
<td>54</td>
<td>40</td>
<td>Blackburn +14</td>
</tr>
<tr>
<td>CBS News/YouGov</td>
<td>10/2 - 10/5</td>
<td>871 LV</td>
<td>--</td>
<td>50</td>
<td>42</td>
<td>Blackburn +8</td>
</tr>
<tr>
<td>FOX News</td>
<td>9/29 - 10/2</td>
<td>666 LV</td>
<td>3.5</td>
<td>48</td>
<td>43</td>
<td>Blackburn +5</td>
</tr>
<tr>
<td>CNN</td>
<td>9/11 - 9/15</td>
<td>723 LV</td>
<td>4.3</td>
<td>45</td>
<td>50</td>
<td>Bredesen +5</td>
</tr>
</tbody>
</table>

All Tennessee Senate - Blackburn vs. Bredesen Polling Data
Questions?