Statistical Science and the Art of Election Polling

Scott Keeter
Director of Survey Research

Prepared for the CNSTAT public seminar
"Reflections on Election Polling and Forecasting from Inside the Boiler Room"

October 19, 2012
skeeter@pewresearch.org
Critics of Polling Were Very Vocal in 2012

UnSkewed Polls.com
a QStarNews website

General Election: Romney vs. Obama

Unskewed Polling Data
Thursday, October 18, 2012 11:06 AM

<table>
<thead>
<tr>
<th>Poll</th>
<th>Date</th>
<th>Sample</th>
<th>Mitt</th>
<th>Paw</th>
<th>Obama</th>
<th>Romney</th>
<th>% Romney</th>
<th>Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poll</td>
<td>8/24 - 9/14</td>
<td>800</td>
<td>44.0</td>
<td>44.2</td>
<td>11.2</td>
<td>44.2</td>
<td>44.0</td>
<td>44.0</td>
</tr>
<tr>
<td>Real Clear/Unskewed</td>
<td>8/11 - 8/14</td>
<td>1500</td>
<td>51.0</td>
<td>46.8</td>
<td>3.1</td>
<td>46.8</td>
<td>51.0</td>
<td>46.8</td>
</tr>
<tr>
<td>Real Clear/ABC News</td>
<td>8/10 - 8/13</td>
<td>805</td>
<td>47.6</td>
<td>46.0</td>
<td>3.5</td>
<td>47.6</td>
<td>46.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Real Clear/CanoePulse</td>
<td>8/18 - 8/21</td>
<td>1360</td>
<td>59.0</td>
<td>35.0</td>
<td>7.0</td>
<td>59.0</td>
<td>35.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Real Clear/YouGov</td>
<td>8/19 - 8/22</td>
<td>600</td>
<td>40.0</td>
<td>43.0</td>
<td>9.0</td>
<td>40.0</td>
<td>43.0</td>
<td>9.0</td>
</tr>
<tr>
<td>REAL Clear/Mel}</td>
<td>9/2 - 9/2</td>
<td>1100</td>
<td>43.0</td>
<td>41.0</td>
<td>6.0</td>
<td>43.0</td>
<td>41.0</td>
<td>6.0</td>
</tr>
<tr>
<td>RealClear</td>
<td>9/2 - 9/2</td>
<td>797</td>
<td>43.0</td>
<td>43.0</td>
<td>3.0</td>
<td>43.0</td>
<td>43.0</td>
<td>3.0</td>
</tr>
<tr>
<td>RealClear</td>
<td>9/24 - 9/27</td>
<td>1112</td>
<td>43.0</td>
<td>41.0</td>
<td>6.0</td>
<td>43.0</td>
<td>41.0</td>
<td>6.0</td>
</tr>
<tr>
<td>RealClear</td>
<td>10/12 - 10/14</td>
<td>1500</td>
<td>46.8</td>
<td>51.0</td>
<td>3.2</td>
<td>46.8</td>
<td>51.0</td>
<td>3.2</td>
</tr>
<tr>
<td>RealClear</td>
<td>10/2 - 10/5</td>
<td>9500</td>
<td>49.0</td>
<td>46.0</td>
<td>1.0</td>
<td>49.0</td>
<td>46.0</td>
<td>1.0</td>
</tr>
<tr>
<td>RealClear</td>
<td>9/28 - 9/30</td>
<td>783</td>
<td>49.0</td>
<td>35.0</td>
<td>6.0</td>
<td>49.0</td>
<td>35.0</td>
<td>6.0</td>
</tr>
<tr>
<td>RealClear</td>
<td>9/26 - 9/29</td>
<td>532</td>
<td>49.0</td>
<td>48.0</td>
<td>1.0</td>
<td>49.0</td>
<td>48.0</td>
<td>1.0</td>
</tr>
<tr>
<td>RealClear</td>
<td>9/24 - 9/25</td>
<td>2652</td>
<td>43.0</td>
<td>45.0</td>
<td>2.0</td>
<td>43.0</td>
<td>45.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

A Tweet in Response to Our Late July Poll
Showing Obama With a Sizeable Lead

David Burge
@lowahawkblog

Say hello to the most incompetent man in America people-press.org/about/scott-ke...

6:30 PM - 2 Aug 12 - Embed this Tweet
But National Polls Have Been Very Accurate

**Candidate Error in National Polling, 1992-2008**

<table>
<thead>
<tr>
<th>Year</th>
<th># of polls</th>
<th>Candidate Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>7</td>
<td>1.1</td>
</tr>
<tr>
<td>1996</td>
<td>9</td>
<td>2.1</td>
</tr>
<tr>
<td>2000</td>
<td>10</td>
<td>1.1</td>
</tr>
<tr>
<td>2004</td>
<td>16</td>
<td>0.9</td>
</tr>
<tr>
<td>2008</td>
<td>19</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Source: NCPP

And State Polls Have Done Nearly As Well

**Candidate Error in State Polling, 2002-2010**

<table>
<thead>
<tr>
<th>Year</th>
<th># of polls</th>
<th>Candidate Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>98</td>
<td>2.3</td>
</tr>
<tr>
<td>2004</td>
<td>198</td>
<td>1.7</td>
</tr>
<tr>
<td>2006</td>
<td>152</td>
<td>2.0</td>
</tr>
<tr>
<td>2008</td>
<td>236</td>
<td>1.8</td>
</tr>
<tr>
<td>2010</td>
<td>202</td>
<td>2.1</td>
</tr>
</tbody>
</table>

Source: NCPP
Thinking about the Challenges to Polls from a Total Survey Error Perspective

Total Survey Error

Errors of Non-Observation

• Coverage Error
  – Cell phone only voters
• Non-response Error
  – Partisanship; engagement
• Sampling Error
Total Survey Error

Errors of Observation

• Questions and questionnaires
  – Determining likely voters and who they will vote for
• Interviewers
• Mode
• Data analysis
  – What to do with the undecided?

Coverage Error
**Telephone Coverage**

Source: National Health Interview Survey

---

**Telephone Coverage**

Source: National Health Interview Survey
Source: Current Population Survey and Pew Research Center surveys

Percent of Adults Ages 18-29 in Landline Samples

Source: Current Population Survey and Pew Research Center surveys
Vote for President
2008 election weekend
likely voter estimates, by sample type

<table>
<thead>
<tr>
<th></th>
<th>McCain</th>
<th>Obama</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual Frame</td>
<td>42</td>
<td>49</td>
</tr>
<tr>
<td>Landline</td>
<td>44</td>
<td>44</td>
</tr>
</tbody>
</table>

Source: Final 2008 pre-election survey by the Pew Research Center

Non-Response Error
Steep Decline in Response Rates Over the Past 15 Years

- Rates calculated using AAPOR's CON2, COOP3 and RR3. Rates are typical for Pew Research surveys conducted in each year.

Demographic Characteristics of Survey Samples are Close to Parameters

<table>
<thead>
<tr>
<th></th>
<th>Pew Research Standard survey</th>
<th>Government surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. citizen</td>
<td>95</td>
<td>92</td>
</tr>
<tr>
<td>Homeowner</td>
<td>63</td>
<td>62</td>
</tr>
<tr>
<td>Lived at current address 5+ yrs</td>
<td>56</td>
<td>59</td>
</tr>
<tr>
<td>Married</td>
<td>50</td>
<td>54</td>
</tr>
<tr>
<td>Children in household</td>
<td>37</td>
<td>37</td>
</tr>
</tbody>
</table>
### Household Financial Circumstances among Responding and Non-Responding Households in a Consumer Database

<table>
<thead>
<tr>
<th></th>
<th>Landline households who...</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Responded</td>
<td>Did not respond</td>
<td></td>
</tr>
<tr>
<td><strong>Net worth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$500,000+</td>
<td>23</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>&lt;$25,000</td>
<td>19</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td><strong>Home value</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$500,000+</td>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>&lt;$25,000</td>
<td>29</td>
<td>26</td>
<td></td>
</tr>
</tbody>
</table>

### Political Characteristics of Responding and Non-Responding Households in a Voter Database

<table>
<thead>
<tr>
<th></th>
<th>Landline households who...</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Responded</td>
<td>Did not respond</td>
<td></td>
</tr>
<tr>
<td><strong>Party registration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Republican</td>
<td>17</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Democrat</td>
<td>23</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>No record of party</td>
<td>54</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Registered to vote</td>
<td>82</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Voted in 2010</td>
<td>54</td>
<td>44</td>
<td></td>
</tr>
</tbody>
</table>
But Sizeable Differences in Civic and Political Engagement

<table>
<thead>
<tr>
<th>In the past year...</th>
<th>Pew Research Standard survey</th>
<th>Government surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volunteered for an organization</td>
<td>55 %</td>
<td>27 %</td>
</tr>
<tr>
<td>Contacted a public official</td>
<td>31 %</td>
<td>10 %</td>
</tr>
<tr>
<td>Talked with neighbors weekly or more</td>
<td>58 %</td>
<td>41 %</td>
</tr>
</tbody>
</table>

Measurement Error: Who’s a Likely Voter
Components of a Likely Voter Index

**Attention/Interest**
- Thought given to election: A lot/Some
- How closely follow election news: Very/Fairly closely
- How much general interest in politics: A great deal
- OR Follow gov’t/public affairs: Most/Some of time

**Past Behavior**
- Voted in previous presidential election: Yes
- How often do you vote: Always/nearly always

**Logistics**
- Ever voted in current precinct: Yes
- Know where to vote: Yes

**Intention**
- How certain to vote: Absolutely certain
- Likelihood of voting (10-pt scale): 9 or 10

---

Why Is It Necessary to Identify Likely Voters?

<table>
<thead>
<tr>
<th>2008 President</th>
<th>Pew election weekend</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Registered</td>
<td>Likely</td>
</tr>
<tr>
<td>Obama</td>
<td>50</td>
<td>49</td>
</tr>
<tr>
<td>McCain</td>
<td>39</td>
<td>42</td>
</tr>
<tr>
<td>Obama margin</td>
<td>+11</td>
<td>+7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2010 House</th>
<th>Pew election weekend</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Registered</td>
<td>Likely</td>
</tr>
<tr>
<td>Rep candidate</td>
<td>43</td>
<td>48</td>
</tr>
<tr>
<td>Dem candidate</td>
<td>44</td>
<td>42</td>
</tr>
<tr>
<td>Rep margin</td>
<td>-1</td>
<td>+6</td>
</tr>
</tbody>
</table>
Share of Two-Party Vote, October 2012

Romney  Obama

Measurement Error:
Asking about Candidate Preference
If the presidential election were being held TODAY, would you vote for...

Analysis:
Dealing with the Undecided

- Split them evenly between the candidates
- Split them proportionally between the candidates
- Ignore them (same as splitting them proportionally)
- Predict their split using their answers to other questions
The “Whiny Leaner” Question

“I have been asked to try to get a choice for president from everyone I interview. Even though you may not have made up your mind, could you make a guess so I will have something to put down? The candidates are....”
Methodological Issues in Conducting Exit Polls

Joe Lenski
October 19, 2012
CNSTAT
Joe Lenski

1. Executive Vice President and co-founder of Edison Research
Edison Research currently conducts all United States exit polling for the National Election Pool (NEP), consisting of ABC, CBS, CNN, FOX, NBC and the Associated Press

- California Recall Election October 2003
- Kentucky and Mississippi Governor Election November 2003
- 23 Democratic Presidential Primaries and Caucuses – January to March 2004
- General Election November 2004
- General Election November 2006
- 40 Presidential Primaries and Caucuses – January to June 2008
- General Election November 2008
- Georgia Senate Run-off – December 2008
- General Election November 2009 – New Jersey and Virginia Governor, and New York City Mayor
- General Election November 2010
- 20 Republican Presidential Primaries & Caucuses – January to April 2012
- General Election November 2012
Warren Mitofsky

- Executive Director of the CBS News Election & Survey Unit 1967-1990; Director of Voter Research & Surveys 1990-1993; President of Mitofsky International 1994-2006
- Developed and tested original methodology for the first exit poll conducted for a U.S. news organization - 1967 Kentucky Governor’s election for CBS News – subsequently conducted more than 3,000 exit polls in the United States, Russia, Mexico, Taiwan and Azerbaijan
Unique aspects of election exit polling

1. Sample universe of locations using past voting data from each location
2. Ask questions to respondents right after they have voted
3. Sample respondents from flow of traffic exiting the polling place
4. Determine age, race & sex of non-respondents (refusals & misses)
Unique aspects of election exit polling

5. Privacy and confidentiality of responses
6. Conduct interviews with a large number of respondents in a single day at a large number of locations
7. Ask questions to discover the reasons behind voters’ choices
8. Immediately transmit results by phone or by handheld device
9. Adjust results to match actual vote returns reported after the polls have closed
1. Sample universe of locations using past voting data from each location

a) Select precincts using a stratified probability sample

b) Collect vote returns for every precincts for the past election used for sample selection

c) Stratify and order precincts by geographic regions and past party vote

d) Select precincts using a random start and interval of selection
1. Sample universe of locations using past voting data from each location
Researching the sample precincts

a) Precinct boundary and name changes
b) Past vote data
c) Voter Registration
d) Polling place location
e) Number of precincts voting at polling place location
f) Poll opening and closing times
g) Ethnic percentages (African-American and Hispanic)
h) Absentee and early voting rules and procedures
2. Ask questions to respondents right after they have voted

3. Sample respondents from flow of traffic exiting the polling place
4. Determine age, race & gender of non-respondents (refusals & misses)

Refusals and Misses Code Sheet:

<table>
<thead>
<tr>
<th>Refusal or Miss (Circle One)</th>
<th>Sex (Circle One)</th>
<th>Age (Circle One)</th>
<th>Race (Circle One)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refusal</td>
<td>Miss</td>
<td>M</td>
<td>18-29</td>
</tr>
<tr>
<td>Miss</td>
<td></td>
<td>F</td>
<td>30-59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>60+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusal</td>
<td>Miss</td>
<td>M</td>
<td>18-29</td>
</tr>
<tr>
<td>Miss</td>
<td></td>
<td>F</td>
<td>30-59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>60+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusal</td>
<td>Miss</td>
<td>M</td>
<td>18-29</td>
</tr>
<tr>
<td>Miss</td>
<td></td>
<td>F</td>
<td>30-59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>60+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusal</td>
<td>Miss</td>
<td>M</td>
<td>18-29</td>
</tr>
<tr>
<td>Miss</td>
<td></td>
<td>F</td>
<td>30-59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>60+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusal</td>
<td>Miss</td>
<td>M</td>
<td>18-29</td>
</tr>
<tr>
<td>Miss</td>
<td></td>
<td>F</td>
<td>30-59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>60+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusal</td>
<td>Miss</td>
<td>M</td>
<td>18-29</td>
</tr>
<tr>
<td>Miss</td>
<td></td>
<td>F</td>
<td>30-59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>60+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. Privacy and confidentiality of responses
6. Conduct interviews with a large number of respondents in a single day at a large number of locations

- Computer system takes in data and delivers the results in real-time to dozens of news organizations around the world
- 876 polling locations covered with an exit poll interviewer in 2010
- More than 91,000 interviews conducted
- More than 60,000 questionnaires processed including over 10,000 telephone surveys of absentee and early voters
- Computations conducted and exit poll analysis delivered for 26 state surveys and one national survey
Hiring and training the interviewers
Network of Interviewers

- Edison has a network of 15,802 trained interviewers across the country who have worked for us for our election and consumer exit polls
- Allows us to cover almost any location in the country
Locations in which Edison Research has conducted commercial exit polls

- Movie Theaters
- Retail Stores
- Shopping Malls
- Concert venues
- Airports
- Sports Arenas
- Restaurants
- Health Clubs
- Museums
- Bars
- Book stores
- Bowling alleys
- Check cashing stores
- College campuses
- Salons
- Veterinarian clinics
- Golf Courses
- Bus Stops
- Subway stations
- Truck Stops
- Grocery stores
- School Buses
- Office Buildings
- Medical Centers
- Convenience stores
- Convention centers
- Cruise ship docks
- Ferry terminals
- Gas stations
- Hotels
- NASCAR races
- Taxi cabs
7. Ask questions to discover the reasons behind voters’ choices

[A] Are you:
1. Male
2. Female

[B] Are you:
1. White
2. Black
3. Hispanic/Latino
4. Asian
5. American Indian
6. Other

[C] In today’s election for U.S. Senate (full term), did you just vote for:
1. Charles Schumer (Dem)
2. Jay Townsend (Rep)
3. Other: Who? ______________
4. Did not vote

[D] In today’s election for U.S. Senate (unexpired term), did you just vote for:
1. Kirsten Gillibrand (Dem)
2. Joseph DioGuardi (Rep)
3. Other: Who? ______________
4. Did not vote

[E] Are you of Hispanic or Latino descent?
1. Yes
2. No

[F] In which age group are you?
1. 18-24
2. 25-29
3. 30-39
4. 40-44
5. 45-49
6. 50-59
7. 60-64
8. 65 or over

[G] In today’s election for governor, did you just vote for:
1. Andrew Cuomo (Dem)
2. Carl Paladino (Rep)
3. Other: Who? ______________
4. Did not vote

[H] What was the last grade of school you completed?
1. No high school diploma
2. High school graduate
3. Some college/assoc. degree
4. College graduate
5. Postgraduate study

[I] Which is closer to your view:
1. Government should do more to solve problems
2. Government is doing too many things better left to businesses and individuals

[J] Do you have any children under 18 living in your household?
1. Yes
2. No

[K] Do you approve or disapprove of the way Barack Obama is handling his job as president?
1. Strongly approve
2. Somewhat approve
3. Somewhat disapprove
4. Strongly disapprove

[L] Which ONE of these four issues is the most important facing the country?
(CHECK ONLY ONE)
1. The war in Afghanistan
2. Health care
3. The economy
4. Illegal immigration

[M] How do you feel about the Tea Party movement?
1. Strongly support
2. Somewhat support
3. Neutral
4. Somewhat oppose
5. Strongly oppose

[N] In the 2008 election for president, did you vote for:
1. Obama (D)
2. McCain (R)
3. Other
4. Didn’t vote

[O] How worried are you about the direction of the nation’s economy in the next year?
1. Very worried
2. Somewhat worried
3. Not too worried
4. Not at all worried

[P] Which candidate do you think is honest and trustworthy?
1. Only Andrew Cuomo
2. Only Carl Paladino
3. Both of them
4. Neither of them

[Q] Does someone in your household belong to a labor union?
1. Yes
2. No

[R] No matter how you voted today, do you usually think of yourself as a:
1. Democrat
2. Republican
3. Independent
4. Something else

[S] On most political matters, do you consider yourself:
1. Liberal
2. Moderate
3. Conservative

[T] 2009 total family income:
1. Under $30,000
2. $30,000 - $49,999
3. $50,000 - $74,999
4. $75,000 - $99,999
5. $100,000 - $199,999
6. $200,000 or more

PLEASE TURN THE QUESTIONNAIRE OVER

New York (G-S-Y1-2010)
8. Transmit results by phone or by handheld device
9. Adjust results to match actual vote returns reported after the polls have closed

- On election night we collect the actual vote returns for every precinct in our exit poll sample.
- We are able to compare the results precinct by precinct to determine “Within Precinct Error” for the exit poll.
- Later on election night we receive the actual vote returns for each county.
- The final exit poll data can then be adjusted to “fact” – the actual vote for each candidate by station and region.
Current issues in exit polling
Current issues in exit polling

1. Increasing proportion of votes being cast before election day
2. Decreasing response rates
3. Differential non-response rates
4. Interviewer-respondent interaction
5. Security of exit poll data before poll closing
1. Increasing proportion of votes being cast before election day

- Many states have extended the ability of voters to cast votes before election day either through in-person early voting or by-mail absentee voting.

- The percentage of voters who cast their vote before election has increased from 16% in 2000 to an estimated 22% in 2004 and it grew to an estimated 32% in 2008 and in 2010.

- Two states – Oregon and Washington - now conduct all voting by mail – ten other states – Arizona, California, Colorado, Florida, Georgia, Nevada, New Mexico, North Carolina, Tennessee and Texas – have more than 45% of votes cast before election day.
1. Increasing proportion of votes being cast before election day

- Voters who cast votes before election day can not be interviewed at the polling place on election day.
- In states that have in-person early voting, Edison Research has experimented with using exit polling methodology at early voting centers – however this adds a sampling factor of time since votes can be cast over a two to three week period.
- In order to include absentee and early voters in our surveys the NEP conducts telephone surveys of absentee/early votes in more than a dozen states with a high proportion of early voting, and merges that data with surveys of voters conducted on election day.
2. Decreasing response rates

- Exit polls in the United States have been experiencing a consistent drop in response rates:
  - 2000 General Election 54%
  - 2004 General Election 53%
  - 2006 General Election 45%
  - 2008 Presidential Primaries 43%
  - 2008 General Election 46%
  - 2010 General Election 45%

- However, response rates for exit polls in other countries where exit polls are a new phenomenon are higher – for example, Edison Research conducted its first exit poll in Puerto Rico for the Democratic Presidential Primary in June 2008 and the response rate was 66%
2. Decreasing response rates

- Other factors effecting response rates:
  - Distance that exit poll interviewer is forced to stand away from polling location exit
  - Weather conditions
  - Multiple exits at polling location
  - Cooperation of local election officials

- In response, the NEP has successfully won legal challenges to 100 foot distance restrictions on exit poll interviewing in six states – Florida, Nevada, Minnesota, New Jersey, Ohio and South Dakota

- The NEP has also introduced smaller questionnaires to help improve response rates
3. Differential non-response rates

There are differential non-response rates by demographic groups. The Edison Research exit polls are able to adjust for non-response by age, race and gender; however, other demographic categories – education, income, political party – can not be identified visually by exit poll interviewers.
4. Interviewer-respondent interactions

- There is evidence of interviewer-respondent interactions that affect both the response rates by demographic group and the accuracy of exit poll estimates.
- Older voters are less likely to complete an exit poll questionnaire handed to them by a younger interviewer.
5. Security of exit poll data before poll closing

- The national news organizations have made a pledge to Congress that they will not report any exit poll results that “characterize” the outcome of the election until all polls in that state have closed.
- In 2004 early exit poll results were leaked to several internet web sites before the polls closed.
- Beginning in 2006 the NEP established a “Quarantine Room” in which a limited number of survey data analysts were allowed to have access to the exit poll data before 5PM.
ELECTION PROJECTIONS: Estimating Election Outcomes

Clyde Tucker
and
Murray Edelman
In 1967, the CBS News Election and Survey Unit, led by Warren Mitofsky, began to develop models of election outcomes using the latest methods for statistical estimation.

These models used actual election returns from a random sample of precincts collected shortly after polls closed and then supplemented with election night county reports.

At first exit polls were used in a few states with dual poll closing times, to account for the part of the state that was still voting. In 1970 CBS pioneered the use of exit polls to analyze the vote. NBC and ABC soon followed. In 1980 NBC used exit polls by themselves to project winners at poll closing. In 1982 CBS and ABC followed.
• In early 1990 ABC, CBS, NBC consolidated their data collection operations under Mitofsky and his colleague, Murray Edelman, who also was responsible for the quality of the data and refining the methods.

• This new organization, Voter Research and Surveys (VRS) included CNN, added AP in 1993, and consolidated the projection system with the News Election System that tabulated the vote.

• VRS became the Voter News Service (VNS), and Mitofsky left and Edelman remained. Fox joined in 1996.
VOTER NEWS SERVICE (VNS)
Precinct Model Estimation

• The estimation follows the method of sampling, described previously.

• All precincts and counties in a state assigned geographic strata codes and party strata codes (based on % Dem in the past race). (Note that the party strata cut points come from ordering the population of all precincts in the state.)

• Past votes (from 3 past races) adjusted for precinct boundary changes available for ratio estimation

• Precinct weight – inverse probability of selection
• Besides collecting results from the precincts after the polls closed, exit polls are conducted in a subsample of the precincts.

• Simple estimates for the geographic and party strata are formed by summing the weighted candidate vote for the precincts in each strata and inflating for any missing precincts.

• A Ratio Party estimator is formed using the same procedure with the best correlated past vote and multiplied by the actual vote in the state. A Ratio Geographic estimator is formed similarly except that the ratios are formed within each stratum and multiplied by the candidate vote in that stratum.
• Variances for simple estimates from exit polls and sample precincts based on weighted squared deviations of current candidate estimates in the precincts from the overall stratum estimate for the candidate

• For the corresponding ratio estimates, those precinct deviations are each subtracted from the ratio used in estimation multiplied by a parallel deviation of the past precinct vote from the past stratum vote

• Standard error created for each stratum--the square root of the sum of squares (SS) divided by the sum of the weights
• The statewide standard error for a candidate -- the square root of the sum (across strata) of each stratum SS weighted by the square of the stratum weight divided by the sum of the stratum precinct weights

• Relying on expert reviews of pre-election polls, a prior estimate made for each election outcome with a variance based on the judgment of the quality of the prior information

• This prior estimate composited with the best precinct estimate using weights inversely proportional to their variances
• To account for the growth in early voting, telephone polls using random-digit dialing (RDD) conducted the weekend before the election in states with large numbers of early voters.

• A composite estimate made of the exit poll with the RDD survey using weights directly proportional to the expected sizes of the election-day vote and the early vote.
After poll closing, actual returns compared to survey results in the exit poll precincts.

Once 10 exit poll precincts overlaid, an estimate of within precinct error (WPE) formed and updated as more actual returns come in.

Once the actual returns overlay the survey data three estimates from the exit poll precincts available—one with precincts with just survey results, one with precincts with just actual returns, and one with a combination of the two.
County Model Estimation

- County estimates based on the vote collected by the AP at over 3,000 counties nationally.

- For each race, counties with minimal data are treated as not reporting. The candidate vote in each other county is inflated by a ratio of the number of precincts in the county to number reporting.
• Candidate estimate by strata—past candidate vote multiplied by ratio of current candidate vote to candidate vote in past race using counties with data

• Proportion of vote for candidate in state—sum of strata estimates divided by estimate of total vote

• An estimate of the error in each county comes from a historical analysis of counties of differing sizes and proportion of precincts reported. The “variance” of the county model estimate is the of sum of the square of the product of each county error and county’s size

• The county geographic model composited with the best precinct model (using weights based on variances) to create an “integrated” model usually with smallest variance
Recent Changes

• In 2003, Edison Media Research (EMR) and Mitofsky International were hired to produce a projection system for the newly formed consortium of networks and AP (the National Election Pool or NEP)

• Instead of providing ratio estimates using only the best correlated past race, ratio estimates using multiple past races are now available
• The field work originally required detailed work with precinct maps to estimate the past vote where boundary changes occurred.

• A change was made to the estimation procedure that made it less dependent on this expensive field work.

• Now the models are not based on weighted precinct vote, but on the average of candidate percentages and the estimates of strata size use a ratio of current vote in the sample precincts to their registration times the registration in the stratum.
• Alternative models based on vote count are also available

• Decision makers have a variety of choices—models with percentages or vote counts, with or without absentees, and with or without remaining precincts with just exit poll data once vote returns become available

• Now that cell-only households have increased to over 30% and approaches 50% when cell-mostly households are included, dual-frame designs (combining landline and cell phone frames) have been used since 2010
One Final Note

• Besides projecting general elections, election projections are made for presidential primaries. The 2008 primaries were particularly exhausting because both parties had contested primaries.

• Projecting primary elections are more difficult than general elections—fewer useful past races exist, order of finish has to be determined when there are 3 or more candidates, caucuses are more complicated than primaries, and primary rules vary more state-to-state than in general elections.