

Do We Know Why Trump Won?

by

Walter Hill

St. Mary's College of Maryland

wwhill@smcm.edu

This draft was prepared for the American Statistical Association, Ann Arbor,
Michigan Chapter meeting of 21 February 2017.

1. BACKGROUND – RECALL ELECTION RESULT

At the end of 2016, Significance magazine noted, “In becoming US President-elect last month, Donald Trump confounded the pollsters, the election modellers and the commentariat, who seemed (almost) certain that Hillary Clinton would win the White House. What went wrong?”¹ Media watchers could be excused for believing that Clinton would become president and the election itself would only be a formality. However, as the evening of 8 November turned into the morning of 9 November it became clear that Donald Trump would become the 45th president. Trump had vaulted over the proclaimed firewall states including North Carolina and Florida. These are states mainstream observers said he could not win. What do we know about the election?

1. 1 SOME FORECAST MODELS

At the outset recall that several forecasts did predict Trump would win, maybe. The journal “PS: Political Science and Politics” has regularly published forecasts of presidential elections. By far, the most popular technique appearing in this publication is multiple regression or a related technique. The range of predictions in this cycle ranged from the forecasts that Democrat would receive between 47.5 percent to 52.7 percent of the two party vote. Of the 9 models, 2 predicted a Trump win. Of those predicting a Republican win, Norpoth was cited in the press. He in fact predicted the fewest votes for the Democrat. His article is “Primary Model Predicts Trump Victory”. The model has the outcome of the New Hampshire primary as one of the independent variables.

Curiously, his forecast was that Clinton would lose the popular vote and hence he logically but incorrectly predicted Trump winning by receiving more popular votes than Clinton. Analogously, the majority of the models correctly predicted that Clinton would receive a majority of the vote, but then incorrectly inferred that she would be elected President. ²

These models exclusively use linear regression. Consider an example from economics. Ray Fair predicted a Trump victory, because the former believed incorrectly that the businessman would win the popular vote. A simplified 2 dimensional version of his model shows the idea. The economist posits voters are myopic and retrospective, making a decision based on economic performance in the last quarter before the election. They vote for or against the incumbent party based on the change in the economy in the 3rd quarter of the election year (some of his models take the last 3 quarters). Other models are variations of the idea of using a regression but employing other variables.

Figure 1.1 about here.
2 Dimensional Ray Fair Model

The other commonly cited correct forecast was by Prof. Allan Lichtman, an American Historian at American University. His model is based on “13 keys” or true/false questions about contemporary conditions. A true key favors the incumbent party. Examples of these questions are: Did the incumbent party have more seats in the House of Representatives than in the previous mid-term elections; Was there no serious contest for the incumbent party nomination; Was the country not in recession during the

campaign? In 2016 six of the “keys” were turned against the incumbent party, hence the Democrats were forecast to lose the election. Litchman indicates that he forecasts the popular vote as he says he correctly called Gore’s popular vote win over Bush in 2000. So despite media claims otherwise, perhaps, I hesitate to say his prediction was correct.³

Curiously, Michael Moore published a scenario in July 2016 which correctly anticipated the outcome. Although his forecast is not scientific, as the variables are not specified before the election, it is worth noting. The tone of his writing is that of a warning. He imagines what he called a Rust Belt Brexit. His observation was that recently Republicans have been able to win statewide elections in the rust belt. Michigan, Ohio, Pennsylvania and Wisconsin have elected Republican governors since 2010. He noted that “all Trump has to do” is hold onto the states won by Romney (in the 2012 election), and pick up several mid-Atlantic and Midwestern states. The scenario, given several months before the election, outlined the Republican’s win.⁴

Another method uses a stock market or futures market. We will discuss the Iowa Electronic market below [perhaps put the Iowa market discussion here]. Another market was available in Canada. The British Colombia lottery was giving Trump 7:2 odds against and Clinton 17/100 (\$1.00 would have paid \$1.17 had Clinton won). Curiously, some British bookies were paying Clinton winners before the election had been called.⁵

2. MACRO-POLL RESULTS & WHAT CAN GO WRONG

Many anticipated a Clinton win. Much of what was expected was based on public opinion data. The vast majority of commentators were convinced that Clinton’s win was inevitable. Trump’s win was so much of a surprise to some that the AAPOR

(American Association of Public Opinion Research) established a special committee to study polling associated with the election. They will issue a report in the spring.

Numerous polls immediately preceding the election estimated the national vote. Consider the Real Clear Politics site. One would conclude that Clinton's advantage over Trump was approximately 2 percent. The common inference from that observation was that Clinton would win the election as the candidate winning the popular vote usually wins the election. However, the outcome of the election is determined by combining the states in which a candidate obtains a plurality. The sum of a weighted score of each of those states (and the District of Columbia) is what matters. The weights are roughly proportional to the population, but above a minimum. Hence it is possible for a candidate to receive the most votes and lose the election. By the way, similar rules appear elsewhere. For example, in tennis it possible for one person to win more games, but lose the match if the scores in the sets are 0-6, 6-4, 6-4.

Figure 2.1 about here
Polling Data for the Major Parties

Table 2.0.1 about here.
Real Clear Politics Polls: Expected Candidate Support

It is clear that the summary of the results of polls indicated a close election but they consistently had Clinton winning. The polls ask what fraction of likely voters will support a given candidate. Clinton did in fact win the popular vote. For decades forecasters argued that the winner of the popular vote typically wins in the Electoral

College. The inference was not true in the last election. The British Columbia lottery is an exception as it specifically defines the criterium for the payoff is the Electoral College result.

There was some concern that current polling practices may not give reliable results. The proliferation of cell phones for example may make it difficult to properly construct a sample frame and obtain a simple random sample. Researchers weigh responses to compensate. The New York Times conducted an interesting experiment. It gave several pollsters the same data and asked if the researchers proposed the same winner. It turns out there were different results.⁶

Table 2.0.2 about here.
Varying Estimates with the same raw data

The differences in spread occur because researchers make different predictions of the likelihood of various subpopulations to vote. The result with the highest support for Trump, for example, assumed the highest turnout for whites. The researcher with the highest estimate for Clinton assumes one of the highest Hispanic turnouts. Assumptions about the demographics in turnout vary by only a few percent, but that size is sufficient to generate spreads that are substantively difference.

2.1 NON-TRANSITIVE PREFERENCES

Polling data showed somewhat non-transitive results. The Hill of 18 May 2016 reports polling data. The Hill is a well known Washington journal reporting on politics.

Its poll showed Clinton beating Sanders, as was then apparent in the primary elections. Further, Sanders beat Trump 46% to 42%. Trump would defeat Clinton by 45% to 42%. This type of intransitive result was frequently reported in mid-spring. This poll had 1,021 respondents and was conducted between 14 and 17 May.⁷

Figures 2.1.1 and 2.1.2 about here
Hypothetical Candidate Position in 1 and 2 Dimensions

A simple model from game theory places voters on a one dimensional scale, say from liberal to conservative. The distribution is assumed to be Gaussian centered at 0 with a spread to be determined. Democratic voters are typically on the left and Republican voters on the right. For simplicity, let us say we can determine which candidate is likely to win in an election by seeing who is closest to the median voter. Given Sanders is to the left of Clinton and both are to the left of the origin, Sanders would not perform better than Clinton against Trump. This result is inconsistent with polling data.

Consider a two dimensional model. In two dimensions, one can easily imagine candidate locations consistent with the polls in which Sanders defeats Trump, and Trump is victorious over Clinton. In the diagram Sanders is closer to the origin than Trump. Further, Clinton would win among Democrats. The fact that one needs two dimensions to create this quasi-intransitive result suggests that a variable based explanation of the election will need at least two variables.

3. LIST OF SPECUTLATIONS ON THE OUTCOME

An incomplete list of some common explanations for the outcome of the election is below.

1. Third party votes for Stein.
2. FBI director Comey announces emails.
3. Russian hacking.
4. Third party votes for Johnson [a loss of Republican votes].
5. White males without college.
6. Anti-Woman vote.
7. Sanders defectors.
8. Desire for a Strongman (Samuel 1) [Alexander Hamilton's view of the President].
9. Clinton's and Trumps campaign in the last weeks.

Of the list above, it is possible to respond to some quantitatively.

1. The hypothesis that Stein's votes prevented Clinton from winning started to appear the evening of the election. This is easy to check. Consider the unrealistic assumption most favorable to the hypothesis that everyone who voted for Stein would have voted Democratic had there been two choices. Such a shift still would not have given Clinton enough votes to win. A Democratic strategist might have worried about this scenario as Stein had polled around 4 percent in October. She told followers two days before the election she hoped to obtain 5 percent of the vote. Stein received 1.46 million votes, or 1.1 percent. Stein did have enough votes to change the outcome in a few states, but in Ohio for example Trump won by 450 thousand votes. Stein had only 46 thousand votes. Similarly in Florida Trump won by 110 thousand votes. Stein only had 64 thousand votes. Stein's not altering the outcome in those two states was sufficient to keep the Republican win.

2. Did FBI Director Comey's announcement, less than 2 weeks before the election, regarding the investigation into Clinton's emails, give the election to Trump? It is not

clear how to test this counterfactual. The Iowa Electronic market clearly shows a response to Comey's announcement. Curiously, from this market, Clinton's position had either come close to recovering, or had in fact recovered, by the time of the election.

Figure 3.1 About here.
Iowa Electronic Market

3. Russian Hacking. This veracity of this hypothesis cannot be answered by statistical data. This story has not died as the New York Times (1 Feb 2017) had an article that one of the agents said to be involved in the hacking, Erovinik, was found mysteriously shot.

4. The third party vote for Johnson. This narrative does not fit into the preferences of either party so it is not so prevalent. The Libertarian candidate was polling as much as 13 percent in the summer. One might interpret the high popularity as an indication of the disapproval of both major party candidates. Johnson received 3.3 percent of the vote. It would be natural that most of the vote that he lost went to the Republicans. Johnson could have been a "spoiler", but he was not.

5. White males without a college education put Trump in the White House.

Exit polls say 71 percent of white males without a college degree voted for Trump. He clearly carried this group. The same poll says 61 percent of white females without a college degree voted for Trump (34 percent for Clinton). He clearly carried this group. In fact, ignoring the education variable, Trump carried white females with 54 percent voting for him. Early in autumn some researchers thought Clinton would easily win by

carrying one swing group, white married females. The CNN exit poll does not give the vote for this group, but does say Clinton barely carried married women by 49-47 percent. Given the substantial non-white vote, we can infer that Clinton lost the white married female vote. So can one really point to a key demographic?

Clinton hoped to replicate Obama's coalition. She called it a "coalition of the ascendant". She could not completely replicate the coalition. For example, there was a reduction in black turnout from the last election. The Times looks at selected southern states because race is listed on one's voter registration form. In Georgia the black share of the electorate dropped to 27.6 from 29.9 percent (2012-2016). In Louisiana the black share dropped to 28.5 from 30.1 percent. In North Carolina, the black share dropped to 27 from 29 percent (preliminary data). However, note the black share of the vote was higher in 2016 than in 2004, possibly the more appropriate comparison point.⁸ Turnout dropped by 14 percent in Detroit, and 8 percent in black wards in Philadelphia. These results represent a reduction of black turnout of 10 percent.⁹

Table 3.1
Selected Exit Poll Data

6. Was there an anti-woman vote? There is some anecdotal evidence to this effect.

An interesting article interviews women who voted for Trump. Susan Chira [NYT 14 Jan 2017, 'Women who voted for Trump in their own words'] interviewed women who voted for Trump. Some women voting for Trump expected him to deliver on jobs, while

others saw themselves voting against Bill Clinton. Several had variations on not wanting Bill in the White House.

7. Sanders supporters were hesitant to vote for Hillary. The dissenters saw themselves as given a difficult choice. Close to the election USA Today reported of Sanders supporters who were likely to vote, 82 percent would vote for Clinton.¹⁰ The vast majority voted Democratic. But the remaining few percent of the electorate (representing 18 percent of the 43 percent of Democrats who voted for Sanders) may have swing close states.

8. Desire for a strong man. This hypothesis has not been picked up by polling data.

9. Campaign. To consider this hypothesis, go back to Michal Moore's scenario about how Trump might win. Moore, writing in July, said the Republican might win the election by first maintaining states Romney won. Then he needs to flip (my word) several rust belt states. This is what happened. Moore notes many people in the rust belt states are discontent with the economy. Prognosticators missed the flipped states, why?

Figure 3.2 about here
Flipped States

Sanders said Clinton never visited Wisconsin after winning the nomination and she visited Michigan late. She “took the bait” by spending time with wealthy people raising money [NPR 8 January 2017]. Over the last few days of the campaign Trump on the other hand campaigned like a marathoner spending time in several rust belt states and correctly predicting that he would win several swing states. The CNN exit poll data said 13 percent of the voters decided in the last week who to vote for. The poll says those

voters went slightly for Trump by a 45-43 margin. The NYT exit poll gives about a 48-41 split for voters deciding in the last day or last week. I suspect in the key rust belt states many who made a decision in the last month decided to vote for Trump. Further, polls within the last week of the election showed the contest narrowing in Colorado, Virginia and Michigan. In a publication of 4 November we see that Clinton no longer had a commanding lead in a number of swing states¹¹ and the race in some states was closing.¹²

4. CONCLUSIONS

1. National level polling results were in fact within expected error bars. The inference of who would win the election was inaccurate because victory is determined not by the popular vote, but by the Electoral College.

2. There was a reduction in the votes for Democrats of non-college educated white voters, not a meltdown. Further, with the exception of those with graduate level education, among whites, all education groups voted for Trump and persons in those groups were more likely to vote Republican than they were in 2012.

3. Clinton lost the white female vote.

4. Part of the explanation of the Democratic loss was there was a decreased turnout among Blacks from 2012. While this is true, the Black turnout was higher than in 2004. The latter might be a better reference point given there was a Black presidential candidate running in 2008 and 2012.

5. Clinton did not bother to campaign extensively in upper Midwestern states. Trump flipped a number of those states.

6. Numerous propositions are not falsifiable (ex. effect of Russian hacking or Comey's statement).

Hence, the conventional wisdom immediately after the election was only partially true. We now know the interpretation of what happened in the election requires several different aspects.

END NOTES

(1) “Significance” (2016), Vol. 13 #6, page 1.

(2) Campbell, James A (2016) ‘Forecasting the 2016 American National Elections’ “PS: Political Science and Politics” Vol. 49, Issue 4, October, pages 649-654. The model by Jerome and Jerome-Speziari one of the few that explicitly forecast and electoral college vote. They predict a Clinton win.

(3) See for example an article on Litchman
https://www.washingtonpost.com/news/the-fix/wp/2016/09/23/trump-is-headed-for-a-win-says-professor-whos-predicted-30-years-of-presidential-outcomes-correctly/?utm_term=.b479a753e26c

(4) Michael Moore - <http://readersupportednews.org/opinion2/277-75/38192-focus-5-reasons-why-trump-will-win>

(5) <http://www.cbc.ca/news/canada/british-columbia/us-election-betting-1.3844770>

(6) Cohn, Nate (2016) “New York Times” ‘We Gave Four Good Pollsters the Same Raw Data. They Had Four Different Results.’ 20 Sept.

(7) <http://thehill.com/blogs/ballot-box/presidential-races/fox-news-poll-bernie-sanders-donald-trump-hillary-clinton>

(8) Democrats won the governor’s contest by 4,800 votes in North Carolina. Obama won in 2008 but not 2012. In 2016: Hillary had 2,189,000 votes, Donald had 2,363,000 votes, and other candidates had 165,000 votes. Clinton lost by 174,000 votes. Exit polls put the black vote at 20%, or 943,000 votes, of which 89% were for the Democrat, 8% Republican. With 90% of the black vote, every 10 black votes gave her 8 votes net. Hence Hillary needed 218,000 more black voters or an increase of 23% of the black vote. To win from this demographic alone she required a non-trivial, and politically unreasonable, increase in the black vote.

(9) <https://www.nytimes.com/2016/12/23/upshot/how-the-obama-coalition-crumbled-leaving-an-opening-for-trump.htm>

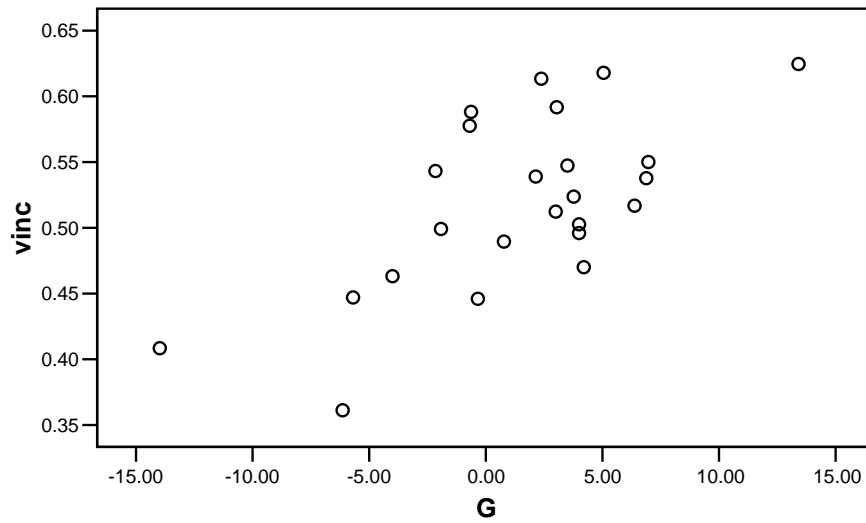
(10) <http://www.usatoday.com/story/news/politics/elections/2016/11/08/bernie-sanders-supporters-election-day/93450782/>

(11) <http://www.vox.com/policy-and-politics/2016/11/4/13502350/hillary-clinton-polls-firewall>

(12) <http://www.ft.com/content/67ff8d88-a1ee-11e6-82c3-4351ce86813f>

A Ray Fair Model

Vote for Incumbent Party and Change in GDP



$r=0.657$, $p<0.001$, $N=24$

$$V = 50.8 + 0.8 \cdot G + U$$

Figure 1.1

Figure 2.1
Polling for major Parties

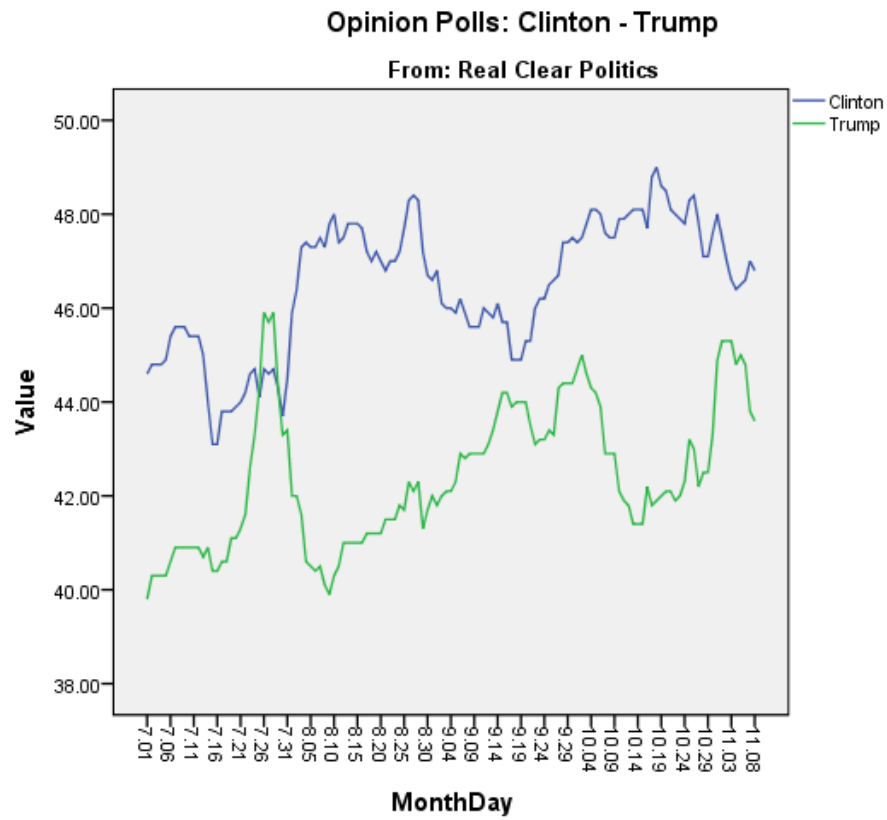


Table 2.0.1

Real Clear Politics Polls close to the election

Poll	Date	N	Type	E	HC	DT	Summary
RCP Average	11/1 - 11/7	--	..	--	46.8	43.6	Clinton +3.2
Bloomberg	11/4 - 11/6	799	LV	3.5	46	43	Clinton +3
IBD/TIPP Tracking	11/4 - 11/7	1107	LV	3.1	43	42	Clinton +1
Economist/YouGov	11/4 - 11/7	3669	LV	--	49	45	Clinton +4
LA Times/USC Tracking	11/1 - 11/7	2935	LV	4.5	44	47	Trump +3
ABC/Wash Post Tracking	11/3 - 11/6	2220	LV	2.5	49	46	Clinton +3
FOX News	11/3 - 11/6	1295	LV	2.5	48	44	Clinton +4
Monmouth	11/3 - 11/6	748	LV	3.6	50	44	Clinton +6
NBC News/Wall St. Jnl	11/3 - 11/5	1282	LV	2.7	48	43	Clinton +5
CBS News	11/2 - 11/6	1426	LV	3.0	47	43	Clinton +4
Reuters/Ipsos	11/2 - 11/6	2196	LV	2.3	44	39	Clinton +5
McClatchy/Marist	11/1 - 11/3	940	LV	3.2	46	44	Clinton +2

http://www.realclearpolitics.com/epolls/2016/president/us/general_election_trump_vs_clinton-5491.html

 Table 2.0.2
 Varying Estimates with the same raw data

Researcher -----	Clinton -----	Trump -----	Spread -----
Franklin: Marquette Law	42	39	Clinton +3
Ruffin: Echelon Insights	39	38	Clinton +1
Omeron et al.: Penn Schoen	42	38	Clinton +4
Corbett-Davis et al.: Stanford	40	41	Trump +1
NYT Upshot/Cienna	41	40	Clinton +1

N=867 Population was Florida, data released 19 Sept. 2016

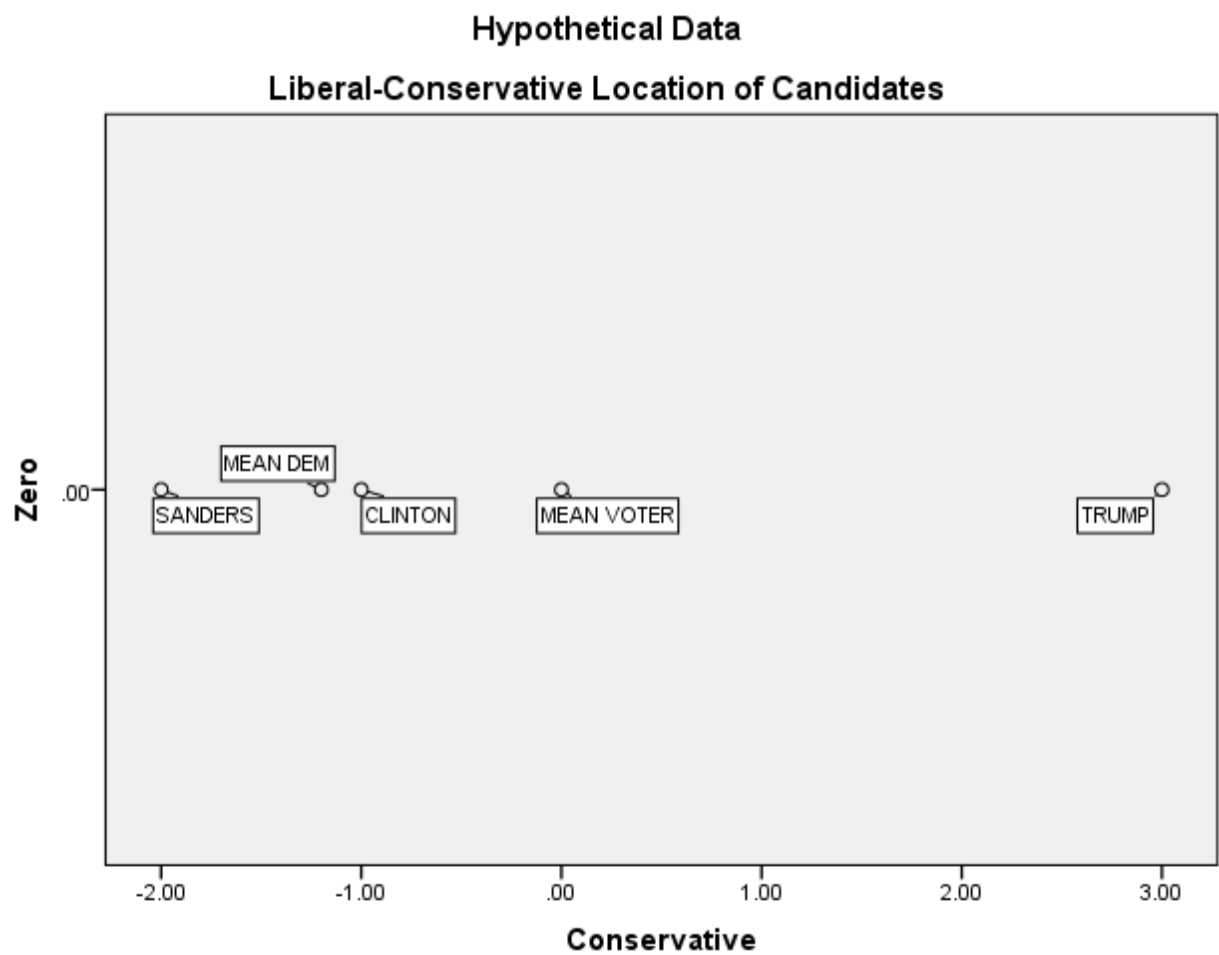


Figure 2.1.1
Candidate positions in 1 dimension

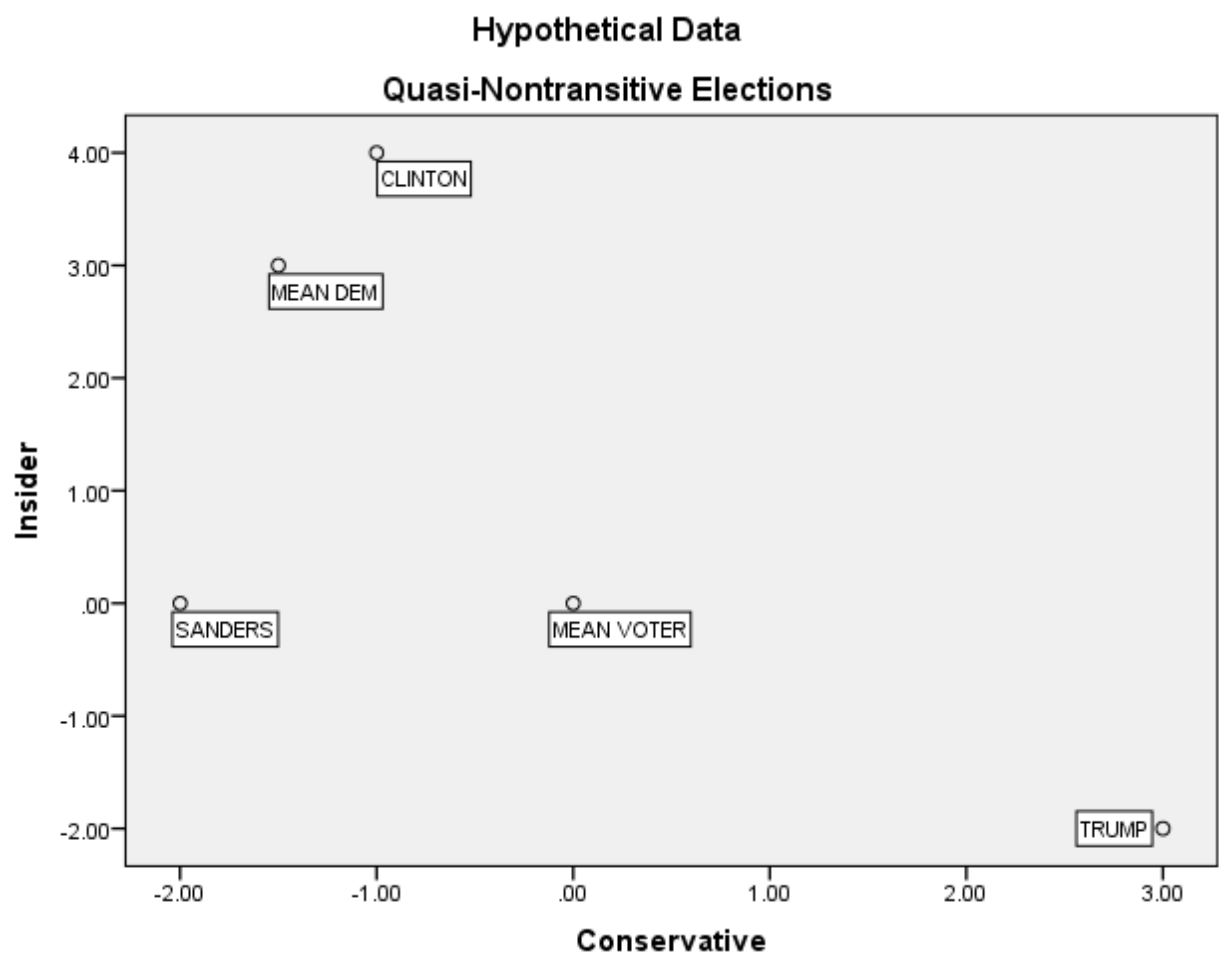
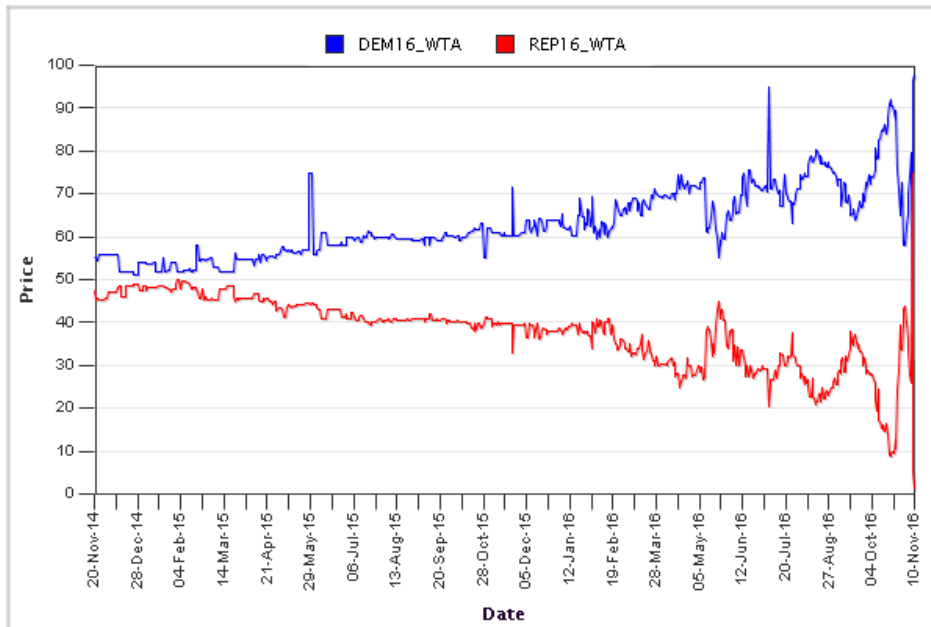


Figure 2.1.2
Hypothetical Candidate positions in 2 dimensions



Sanders supporters disrupt Nevada convention – Sat 14 May.
 Trump University revelations + Albuquerque Trump Demonstrations – About 1 June
 California Primary – Tues. 7 June.
 Republican Convention - Monday-Thursday 18-21 July.
 Democratic Convention - Monday-Thursday 25-28 July.
 First Debate – Monday 26 Sept.
 Trump’s Lewd “Access Hollywood” Tape – Fri 7 Oct.
 Second Debate – Sunday 9 Oct.
 Third Debate – Wednesday 19 Oct
 FBI Director announces discovery of emails – Friday 28 Oct.
 Election - Tuesday 8 Nov.

Figure 3.1
 Iowa Stock Market – Winner Take All Security

Table 3.1
Selected Exit Poll Variables

Selected Exit Poll Variables

Party	% vote Romney Obama			% voters Clinton Trump		
Democrat	38	7	92	36	89	8
Republican	32	93	6	33	8	88
Independent	29	45	50	31	42	46

Ideology						
Liberal	25	11	86	26	84	10
Moderate	41	41	56	39	52	40
Conservative	35	82	17	35	16	81

Ethnic						
White	72	59	39	71	37	57
Black	13	6	93	12	89	8
Hispanic	10	27	71	11	66	28
Asian	3	26	73	4	65	37

Sex/ethnicity						
White men	34	62	35	34	31	62
White women	38	56	42	37	43	52
Black men	5	11	87	5	82	13
Black women	8	3	96	7	94	4
Latino Men	-	-	-	5	63	32
Latina Women	-	-	-	6	69	25

Age						
18-29 yr	19	37	60	19	55	36
30-44 yr	27	45	52	25	51	41
34-64 yr	38	51	47	40	44	52
>=65 yr	16	56	44	16	45	52

Education						
Not high school	3	35	64	-	-	-
High school grad	21	48	51	-	-	-
High school or less	-	-	-	18	46	51
Some college	29	48	49	32	43	51

College grad	29 51 47	32 49 44
Postgrad	18 42 55	18 58 37

Education Among Whites By Sex

White college grad women	20 51 44
White non-college women	17 34 61
White college grad men	17 39 53
White non-college grad men	16 23 71
Non-white	29 74 21

Religion

Protestant	53 57 42	52 39 56
White protestant	39 69 30	- - -
White evangelical	26 78 21	- - -
Catholic	25 48 50	23 46 50
Jewish	2 30 69	3 71 23

Marital Status by Gender

Married Men	29 60 38	29 38 57
Married Women	31 53 46	30 49 47
Unmarried Men	18 40 56	18 46 44
Unmarried Women	23 31 67	23 63 32

Family income

<\$50,000	21 38 60	36 53 41
\$50k-\$100k	59 52 46	30 46 49
>\$100,000	28 54 44	44 47 48

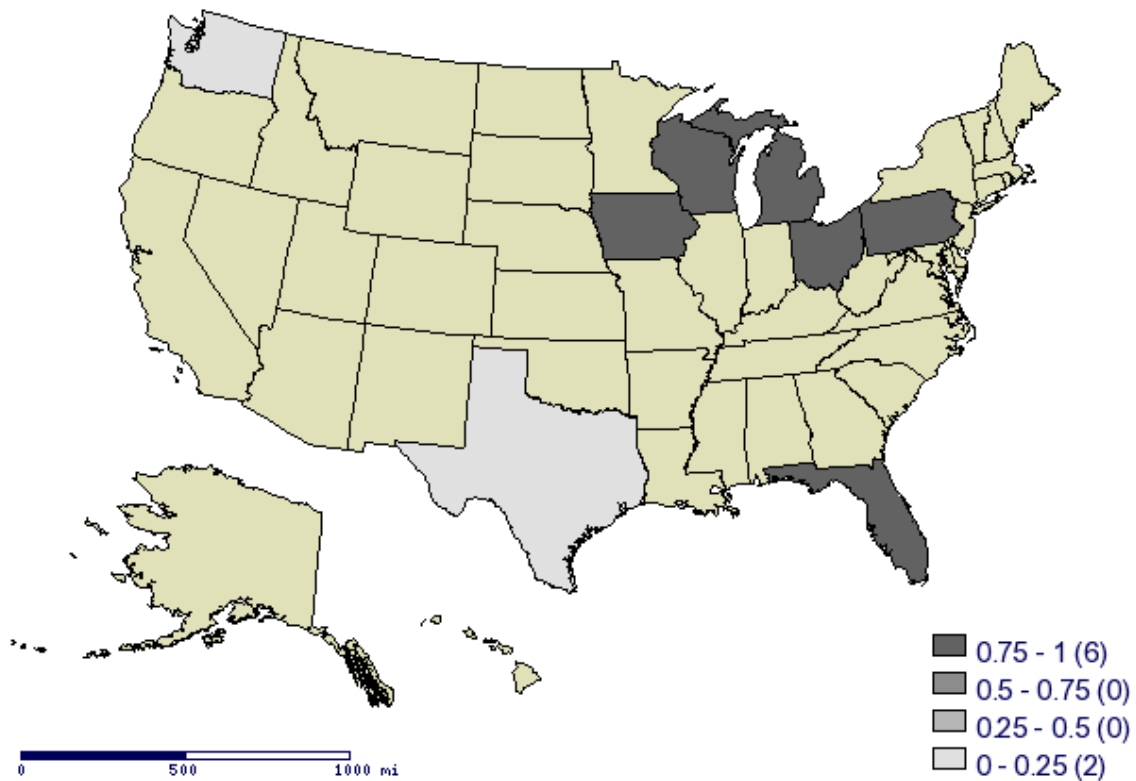
When Did You Decide?

Last few days	9 50 45	8 43 43
Last week	- - -	5 41 49
In October	11 49 4	12 37 51
In September	9 45 53	12 46 48
Before September	69 53 46	60 52 45

2016 data from CNN

2012 data from Edison Research & Schier Box-Steffensmeier

Majority Party in Presidential Vote Switched From 2012 to 2016



Dark states switched from Obama to Trump (Democrat to Republican).
Grey states have at least one faithless elector in 2016.
No states switched from Romney to Clinton (Republican to Democrat).

FIGURE 3.2
Flipped States