Are Democrats Losing Union Rank And File

Eric Oganesian

James Ebersole

Sri Seshadri

Karol Jarzabek  
Northwestern University

Are Democrats Losing Union Rank And File Union Members

**Introduction**

We are an independent political consulting firm headquartered in Chicago with field offices in Washington, DC and Indianapolis, IN. A 527 organization has hired us to help determine what caused the battleground states of Pennsylvania, Ohio, Michigan and Wisconsin to flip from Barack Obama in 2012 to Donald Trump in 2016. To fulfill our client’s directive we are conducting a phone poll of the 25 counties in the four aforementioned states which had at least a 20% shift from supporting the Democratic presidential ticket in 2012 to the Republican presidential ticket in 2016 (Fahey & Wells, n.d). Our initial hypothesis is that trade union affiliation is positively correlated to voters switching from Democrats in 2012 to Republicans in 2016. Our null hypothesis is that there is not a positive correlation between trade union affiliation and a voter's likelihood to have switch from Obama in 2012 to Trump in 2016.

**Sample Population**

The sample for our survey will consist of United States citizens over the age of 22 years old who voted in at least the past two presidential election cycles in counties which swung from Obama in 2012 to Trump in 2016 by at least 20% in the politically important battleground states of Pennsylvania, Michigan, Ohio and Wisconsin. The data of who voted in these counties in both the 2012 and 2016 elections will be procured via NationBuilder at an estimated cost of $1,000 which has already been approved by the client. The NationBuilder data also includes demographic information such as race, gender and age which will be used to formulate a stratified random sample. The counties which switched by at least 20% can be found below in Table 1.

A total of 747,352 ballots were cast in the 25 counties in 2016, using that number as our best estimate as to how many voted in 2012 and 2016 we will need a sample size of 2219 for an expected correlation of 0.8 with a margin of error of 0.03 at a 95% confidence interval is calculated using a 2 stage approximation of sample size based on the desired width of confidence interval as described in (Bonett & Wright, 2000).The algorithm described in (Bonett & Wright, 2000) is coded in open source R and can be referenced in (Weaver & Koopman, 2014). Based off a 9% sample response rate we will need to reach out to 24,409 members of our target population (Keeter, Hatley, Kennedy, & Lau, 2017).

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1: Counties That Flipped By At Least 20%** | | | |
| **Pennsylvania** | **Ohio** | **Michigan** | **Wisconsin** |
| Luzerne | Ashtabula, Trumbull, Ottawa, Sandusky, Loraine, | Monroe, Manistee, Lake, Shiawassee, Gogebic | Sawyer, Lincoln, Forest, Pepin, Buffalo, Trempealeau, Jackson, Juneau, Adams, Marquette, Crawford, Richland, Grant, Lafayette |

**Potential Errors and Design Elements**

We foresee under coverage of union members, misreporting by participants, and non-response as the greatest potential sources of error in this survey.

For the concerns of coverage error, our use of a dependable third party vendor for rich data sets should restrict any major data loss by under coverage by selecting only eligible units (Groves et al., 2009). A source of coverage error could arise from picking a sample population where union members are under represented. By cross referencing county level electoral results with state level data of union membership from the U.S. Bureau of Labor Statistics (“Table 5. Union affiliation of employed wage and salary workers by state,” n.d.), we’ve validated our choice of target counties for the study. By randomizing and ensuring equal chance of selection for every voter record, we’ve adjusted for the possibility of sampling error.

Our intended sample size was calculated taking non-response into account, but there is still the possibility for high levels of non-response with a public saturated by public polling. In the design of our survey, we will employ pretests, interviewer training, interviewer matching, recontacting, and calling at particular times of day in an attempt to decrease our level of non-response. Interviewer shifts will be built around a Sunday – Thursday schedule between 6-9 p.m. and will use these same time frames for recontact attempts. Interviewers will be trained over a two day period and conduct field pretests on a scale with the survey to increase their skills and understanding of the survey while also providing feedback for script or reporting issues. The training on the script and pretests will include emphasis on high-quality data over efficiency, motivational roles for interviewers, and coding and script reviews (Groves et al., 2009).

By setting up a computer-based reporting system, using mostly closed-answer questions, and coding with the Standard Occupational Classification (SOC) and North American Industry Classification System (NAICS) for occupation and employer, we will reduce processing error to a negligible rate (Groves et al., 2009).

Additionally, we’ve included a question asking whether a respondent has voted for President in the 2012 and 2016 elections. Given that we already have that data through a third party, we can use them to track discrepancies in responses among particular groups. We intend to use imputation by subgroups to adjust for nonresponse and measurement error in the post-survey stage (Groves et al., 2009).

## Methods Of Analysis

There are four main methods that will be used to analyze the data collected which are detailed below:

### **Exploratory / Visual Data Analysis (EDA / VDA):**

Our data analysis will begin with exploratory and visual analysis to quickly identify anomalies in the data and to ensure all responses to the survey questions formatted and coded correctly. Data will also be organized in way that allows for easy scatter plots between union and non-union voters that switched the party the voted for in the 2012 and 2016 presidential elections. The visual analysis will help to determine if there is a correlation and its strength if one exists Parallel coordinate plots will help to visualize which issues were important for voters who voted across party lines.

#### **Hypothesis Testing:**

We will statistically test the null hypothesis that there is not a correlation between voters who are union members and those who switched their vote from Obama in 2012 to Trump in 2016. If the results show there is a correlation then we will then be able to reject the null hypothesis and therefore accept the hypothesis that there is a correlation.

#### **Regression and Algorithms:**

A logistical regression model and classification algorithms will be created to determine what policy issues or other factors such as demographic information were likely predict switching from Obama in 2012 to Trump in 2016.

**Conclusion**

The results of our research will have a far reaching impact on how our client identifies and does outreach to voters in swing states and particularly union members. If the hypothesis proves true that there is a correlation between union membership and voters moving towards the Republican Party at a national level then we will likely see both major national political parties adjust their platforms to further court this formerly secure Democrat voting bloc.

References

Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating pearson, kendall and spearman correlations. *Psychometrika, 65*(1), 23-28. doi:10.1007/BF02294183

Fahey, M., & Wells, N. (n.d.). *The Places that flipped and gave the country to Trump*. Retrieved from CNBC: http://www.cnbc.com/heres-a-map-of-the-us-counties-that-flipped-to- Trump-from-democrats/

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2nd ed.). Hoboken, New Jersey: John Wiley & Sons.

Keeter, S., Hatley, N., Kennedy, C., & Lau, A. (2017, May 15). *What low response rates mean for telephone surveys.* Retrieved from Pew Research Center: <http://www.pewresearch.org/2017/05/15/what-low-response-rates-mean-for-telephone-surveys>

Bureau of Labor Statistics. (n.d.). *Table 5. Union affiliation of employed wage and salary workers by state*. Retrieved from <https://www.bls.gov/news.release/union2.t05.htm>

Weaver, B., & Koopman, R. (2014). An SPSS Macro to Compute Confidence Intervals for Pearson’s Correlation. *The Quantitative Methods for Psychology, 10*(1), 29-39. Doi:1 0.20982/tqmp.10.1.p029

Appendix A

Sample Size

|  |
| --- |
| Table A1: R Code for Sample Size |
| knitr::opts\_chunk$set(echo = TRUE) sizeCIcorr <- **function**(alpha, corr, CI2w) {  # Computes sample size required to estimate a correlation with desired precision  # Args:  # alpha: alpha level for 1-alpha confidence  # corr: planning value of correlation  # CI2w: desired confidence interval width  # Returns:  # required sample size  z <- qnorm(1 - alpha/2)  n1 <- ceiling(4\*(1 - corr^2)^2\*(z/CI2w)^2 + 3)  zr <- log((1 + corr)/(1 - corr))/2  se <- sqrt(1/(n1 - 3))  LL0 <- zr - z\*se  UL0 <- zr + z\*se  LL <- (exp(2\*LL0) - 1)/(exp(2\*LL0) + 1)   UL <- (exp(2\*UL0) - 1)/(exp(2\*UL0) + 1)  N <- ceiling((n1 - 3)\*((UL - LL)/CI2w)^2 + 3)  **return**(N) } |

|  |
| --- |
| Table A2: Correlation to Required Sample Size |
| cor SampleSize 0.0 17074 0.1 16734 0.2 15736 0.3 14140 0.4 12049 0.5 9607 0.6 6997 0.7 4445 0.8 2219 0.9 624 1.0 NaN |

|  |
| --- |
| Table A3: Sample Size to Correlation Chart |
|  |

Appendix B

Interviewer Script:

*Notes:* The script will have conditional formatting on computers that will lead the interviewer through the script depending on entered answers. For open-ended questions, short answer spaces will be available to interviewers that will be coded by international standards. Each question in the interviewer prompt will have an option to input “refusal” if the participant is unwilling to answer any of the questions scripted for the interviewer.

* Introduction:
  + “Hi my name is \_\_\_\_\_\_ and I’m calling from Northwestern University. We are currently conducting a survey into the results of the recent 2016 Presidential election. We are not selling anything, simply conducting a survey for academic use where all responses are absolutely confidential.”
* Identify individual:
  + “To start off with, am I speaking with \_\_\_\_\_\_\_?” > “Thank you, is \_\_\_\_\_\_ home?” > “When would be a good time to call back for them?”
  + **OR** “It’s great to speak with you, I hope you’re well today. Do you have 5 minutes to complete a short survey?” > “Thank you. I want to let you know that this call may be recorded for training and survey effectiveness, is this okay?” > “Great, let’s begin!”
* Union membership:
  + “Are you a member of any kind of trade or labor union?” > “Could you tell me which Union membership you hold?”
  + **OR** “Have you ever been a member of a Union?”
* Occupation & Industry:
  + “Could you tell me who your current employer is?”
  + “Could you tell me what your current occupation is?”
* “I’m now going to transition briefly into some questions about elections. While some you may not be able to remember certain answers, please try to be as accurate as possible.”
* 2016 Presidential:
  + “As a simple yes or no, did you vote in the 2016 Presidential election?”
  + “What would you say was the issue or issue area that most influenced your decision in the 2016 Presidential election?”
  + “Are there any other issues or issue areas that were of similar importance to that decision?”
  + “Would you be willing to tell me who you voted for President in 2016?”
* 2012 Presidential:
  + “As a simple yes or no, did you vote in the 2012 Presidential election?
  + “What would you say was the issue or issue area that most influenced your decision in the 2012 Presidential election?”
  + “Are there any other issues or issue areas that were of similar importance to that decision?”
  + “Would you be willing to tell me who you voted for President in 2012?”
* Conclusion: “Thank you so much for your participation and your time. I hope you have a nice evening!”