The Accuracy of Microtargeted Policy Positions

Kyle Endres, *Duke University*

ABSTRACT Identifying voters who share policy positions with the candidates has become an important component of modern political campaigns as they rely on microtargeted estimates to guide targeting decisions. Using survey data and microtargeted estimates from the 2012 election, I conduct one of the first independent examinations of the accuracy of microtargeting. The estimates are the most accurate in Florida, a state that requests information on the race of voters when they register and has party registration. The estimates are less accurate in the other battleground states that do not collect as much information. The accuracy rates range from 36% to 82% depending on the issue and state.

Data and technological advancements have improved the ability of campaigns to contact the right voters with the right issues. Journalistic accounts of Big Data and modern campaigns have touted the accuracy of the information campaigns possess about the American citizenry (e.g. Issenberg 2013). Campaigns use this data to more efficiently target the right voters with the right message(s) (Nickerson and Rogers 2014). Correctly identifying the voters who should be the most responsive to the campaign and correctly identifying areas of agreement between the candidate and these voters should enhance the campaign's ability to effectively mobilize its supporters and persuade swing voters. The use of Big Data to estimate the policy positions and other characteristics of individual voters is commonly referred to as microtargeting.

Microtargeting involves two stages. First, large surveys are conducted to develop statistical models that are used to predict attributes about the voters in a state such as how likely individuals are: to vote, to support each presidential candidate, and to support specific policy positions. These models are then used to estimate these same attributes for all of the individuals listed in a state's voter file (see Panagopoulos 2017, 119). It remains an open question how detailed of a picture microtargeting provides of each voter. Media stories from the 2012 election purported that the "campaigns know you better than you know yourself" (Brennan 2012). However, past studies have yielded mixed results in regards to the accuracy of the microtargeted estimates of the race, religion, and marital status of individual voters, suggesting that the parties and their data vendors may have little to offer when it comes to identifying persuadable voters and where they stand on the issues (Hersh 2015). Using survey data and microtargeted estimates of the policy positions of registered voters in three battleground states from the 2012 election, I evaluate the accuracy of the estimates compared to the voter's self-reported positions.

Research on presidential campaigns has often focused on the policy issues that the candidates should emphasize in their outreach to voters (e.g. Petrocik 1996; Vavreck 2009). Modern campaigns rely on microtargeting to execute their individual level contact strategies. Armed with information on which individuals are the most likely to agree (or oppose) their position on any number of issues should allow campaigns to more effectively execute issue-based mobilization and persuasion strategies. The success of policy based appeals is dependent on the campaign's ability to contact voters who share their candidate's position on the issues (Hillygus and Shields 2008). Thus, correctly identifying where individual voters stand on the issues is critical and warrants examination.

BACKGROUND & DATA

The survey data used for this project are from 2012 Cooperative Congressional Election Study (CCES), conducted by YouGov. The CCES is a national survey with sufficiently large sample sizes in many states to investigate political preferences separately in each state. The CCES consists of both a pre-election wave, which was fielded during October of 2012 and a post-election wave. The Republican National Committee (RNC) provided separate voter files for Colorado, Florida, and Virginia—key battleground states in the 2012 election.[1] The RNC voter files are expanded versions of the voter records maintained by the Colorado Department of State, Florida Department of State, and the Virginia Department of Elections, respectively. The voter files were made available by the RNC to Republican candidates including the 2012 presidential candidate, Mitt Romney.

YouGov regularly matches voter files with its surveys. In fact, the 2008 and 2012 CCES were both matched with voter files for every state. YouGov was provided with the RNC voter files that included the microtargeted estimates and matched them with the survey data using the personal identifying information that it collects but does not make available to researchers.[2] A total of

Kyle Endres *is a postdoctoral associate at the Duke Initiative on Survey Methodology (DISM) at Duke University. He may be reached at kyle.endres@gmail.com.*

doi:10.1017/S1049096516001645

4,095 voters were successfully matched to survey respondents in the CCES. Matching the RNC voter files with the CCES survey data allows me to directly compare the self-reported policy positions from the pre-election wave with microtargeted estimates for the issues that were included in the 2012 CCES.

POLICY ISSUES IN THE 2012 VOTER FILES

A common narrative is that political campaigns estimate where the voters stand on numerous policy issues. In reality, the voter files from 2012 included microtargeted estimates on a limited number of issues. The issues estimated and the format of the estimates differed between the states. The Colorado and Virginia voter files included microtargeted estimates on three issues: the economy, Obamacare, and social conservatism. The Florida file included: abortion, the economy, gun control, Obamacare, jobs, and debt. I test the accuracy of abortion/social conservatism, the economy, gun control, and Obamacare using survey items in the CCES. Issues without a suitable comparison item are not examined. The method and models used to estimate the positions of each individual who is registered to vote is proprietary information and has not been made available to the author. The process behind microtargeting, however, is discussed in great detail elsewhere (see Hersh 2015; Nickerson and Rogers 2014).

The Colorado and Virginia voter files included microtargeted estimates on three issues: the economy, Obamacare, and social conservatism. The Florida file included: abortion, the economy, gun control, Obamacare, jobs, and debt.

The microtargeted estimates in the Colorado and Virginia files were collapsed into categorical variables, with three categories for each issue: a pro-Republican category, a neutral category, and a pro-Democratic category. The estimates in Florida were provided to the author as scale variables. For Colorado and Virginia, I evaluated the accuracy of the pro-Republican category and the pro-Democratic category. The middle category is not examined. For Florida, I collapsed the variables into similar categories and tested the accuracy of the most pro-Republican estimates and the most pro-Democratic estimates. Looking at the pro-Republican estimates is appropriate since the Republicans should only emphasize issues in which an individual voter is expected to agree with them. Conversely, examining the pro-Democratic estimates hint at the percentage of voters the Republicans should have contacted, but were erroneously classified as opposing the Republican position by the microtargeted estimates.

I compared each of the microtargeted estimates to the appropriate variable from the CCES. Abortion, gun control, and Obamacare were all included on the CCES. The estimates were coded as correct if the self-reported position is the same as the estimate. For the economy, I compared the estimate to the respondent's assessment of the direction of the economy over the past year.[3] Indicating the economy had improved during the year prior to the election was coded as the pro-Democratic position and reporting that the nation's economy had worsened over the past year was coded as the pro-Republican position.

Not all of the voters who were matched to the survey data had microtargeted estimates of their policy positions. Table 1 lists the distribution for each variable and includes the percentage of voters who did not have microtargeted estimates in that state. In my sample, 85% of the voters in Colorado had microtargeted estimates, 91% of the voters in Florida had microtargeted estimates, and 82% in Virginia included microtargeted estimates. For each of the policy issues in Colorado and Virginia, larger percentages were estimated to have a pro-Republican position on the issue than were projected to have a pro-Democratic position.

EXPECTATIONS

There are a few noteworthy differences between the Colorado, Florida, and Virginia voter files that may affect the accuracy of the estimates. Colorado and Florida both have party registration, which is included in the files. The Virginia voter file does not include party registration as the state has open primaries and does not give individuals the opportunity to officially register with a political party, but does include information on which party's primary (if any) the voter cast a ballot in previous elections. In addition to information on a registered voter's partisan affiliation, Florida is one of eight states that asks individuals to indicate their race when they register to vote.[4] Partisanship and race of the voter are two of the most valuable data points for campaigns (Hersh 2015). The accuracy of the estimates should reflect the variations in the voter files. Florida estimates should be the most accurate since it collects both party registration and race. Virginia estimates should be the least accurate since its voter files do not include either party registration or race.

Table 1

Distribution of Variables

	Abortion	Economy	Gun Control	Obamacare
Colorado (n=701)				
Pro-GOP	41%	28%	-	37%
Middle	14%	33%	-	25%
Pro-Dem	30%	25%	-	23%
No MT	15%	15%	100%	15%
Florida (n=2,396)				
Pro-GOP	23%	33%	23%	22%
Middle	44%	26%	36%	47%
Pro-Dem	24%	32%	32%	22%
No MT	9%	9%	9%	9%
Virginia (n=998)				
Pro-GOP	40%	32%	-	40%
Middle	20%	28%	-	23%
Pro-Dem	22%	21%	-	19%
No MT	18%	18%	100%	18%

Cells display the percentage of individuals estimated to be in each category for the 4,095 voters who were matched between the voter files and the 2012 CCES.

https://doi.org/10.1017/S1049096516001645 Published online by Cambridge University Press

RESULTS

The Florida estimates were the most accurate across all issues. The comparisons are presented separately for the voters estimated to have a pro-Republican position and for the voters estimated to have a pro-Democratic position in table 2. The combined results are also included in table 2. Voters in Florida who were estimated to have a pro-Republican position self-reported a pro-Republican position 75% of the time for abortion, 77% for the economy, and 81% for both gun control and Obamacare. Based on these percentages, when Republican candidates in Florida contacted voters on one of these issues, the recipient of the campaign message agreed with them most of the time and disagreed with the GOP position between 19% and 25% of the time. These numbers are comparable for the voters estimated to hold pro-Democratic positions with the exception of the economy. Only 59% of Florida voters who were estimated to hold a pro-Democratic evaluation of the economy reported that they believed the economy had improved over the past year when asked the month prior to the election. If a campaign relied on these estimates to determine whom to contact on each issue, they likely missed many voters who may have been receptive to an economic message from Republican candidates.

The microtargeted estimates were less accurate in Colorado and Virginia. The estimates on the economy were the least reliable. Pro-Republican economic estimates were accurate 50% of the time in Colorado and 46% of the time in Virginia. And pro-Democratic estimates were accurate 36% of the time in Colorado and 43% of the time in Virginia. The abortion and Obamacare estimates more closely adhered to the voter's self-reported positions. In Colorado, the pro-Republican estimates had an accuracy rate of 62% for abortion and 70% for Obamacare. In Virginia, the pro-Republican estimates for abortion had a similar accuracy rate at 60% and the Obamacare estimates had an accuracy rate of 51%. Almost half of the voters in Virginia who were estimated to oppose Obamacare indicated they supported it in the CCES. The pro-Democratic estimates for abortion and Obamacare were more accurate in Colorado and Virginia than the pro-Republican estimates. In Colorado they matched the self-reported positions approximately 75% of the time. Sixty-one percent of the abortion estimates were correct in Virginia and 65% of the Obamacare estimates were correct.

Table 2

Accuracy Rates by Issue and State

	Abortion	Economy	Gun Control	Obamacare
Colorado				
Pro-GOP	62%	50%	-	70%
Pro-Dem	76%	36%	-	75%
Combined	68%	43%	-	72%
Florida				
Pro-GOP	75%	77%	81%	81%
Pro-Dem	73%	59%	72%	82%
Combined	74%	68%	76%	82%
Virginia				
Pro-GOP	60%	46%	-	51%
Pro-Dem	61%	43%	-	65%
Combined	60%	45%	-	55%

Cells display the percentage of individuals with a self-reported position in the 2012 CCES that matches their estimated position in the 2012 Republican voter files.

CONCLUSION

Modern campaigns do have the ability to identify which voters support and oppose their positions on the issues central to their candidacy, but this ability varies based on both the amount of information states collect (and include in their voter files) and on the issue. As the use of microtargeted data becomes more widespread in this and future elections, the implications for existing theories on campaign effects and voter persuasion should be reexamined in this context. Campaigns that rely on microtargeted estimates to determine which voters to contact on specific issues are getting it right much of the time, but are far from perfect. The risks of targeting voters with the wrong issue should be weighed against the rewards from contacting other voters with the right issue. However, the effects of issue based partisan appeals—both when the campaign correctly and incorrectly identifies an individual's position—need to be investigated in greater detail.

Modern campaigns do have the ability to identify which voters support and oppose their positions on the issues central to their candidacy, but this ability varies based on both the amount of information states collect (and include in their voter files) and on the issue.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1049096516001645.* ■

NOTES

1. Other battleground states include: IA, MI, NC, NH, NV, OH, PA, WI (Sides and Vavreck 2013).
2. The procedure used for this study is comparable to the matching performed for Ansolabehere and Hersh's voter validation study using the 2008 CCES (See Ansolabehere and Hersh 2012). The main difference is the Republican Party provided the voter files for this study, while Catalist, a commercial vendor that works closely with the Democratic Party provided the voter files for the 2008 and 2012 validated CCES studies.
3. Exact question wording and coding is available in the online appendix.
4. AL, FL, GA, LA, NC, PA, SC, and TN collect racial or ethnic information when individuals register to vote.

REFERENCES

Ansolabehere, Stephen and Eitan Hersh. 2012. "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate." *Political Analysis* 20 (4): 437–59.

Brennan, Allison. 2012. "Microtargeting: How Campaigns Know You Better Than You Know Yourself." CNN. http://www.cnn.com/2012/11/05/politics/voters-microtargeting/

https://doi.org/10.1017/S1049096516001645 Published online by Cambridge University Press

Hersh, Eitan D. 2015. *Hacking the Electorate: How Campaigns Perceive Voters*. New York: Cambridge University Press.

Hillygus, D. Sunshine and Todd G. Shields. 2008. *The Persuadable Voter: Wedge Issues in Presidential Campaigns*. Princeton, NJ: Princeton University Press.

Issenberg, Sasha. 2013. *The Victory Lab*. New York: Broadway Books.

Nickerson, David W. and Todd Rogers. 2014. "Political Campaigns and Big Data." *Journal of Economic Perspectives* 28 (2): 51–73.

Panagopoulos, Costas. 2017. *Political Campaigns: Concepts, Context, and Consequences*. New York, NY: Oxford University Press.

Petrocik, John R. 1996. "Issue Ownership in Presidential Elections, with a 1980 Case Study." *American Journal of Political Science* 40 (3): 825–50.

Sides, John and Lynn Vavreck. 2013. *The Gamble: Choice and Chance in the 2012 Presidential Election*. Princeton: Princeton University Press.

Vavreck, Lynn. 2009. *The Message Matters: The Economy and Presidential Campaigns*. Princeton, NJ: Princeton University Press.

* The URL to access Supplementary Material for this article has been corrected since the original publication. An Erratum detailing this change was also published (DOI: 10.1017/S1049096516001645).