

# Predicting State Presidential Election Results with National Tracking Polls and MRP

Chad P. Kiewiet de Jonge<sup>\*</sup>

Gary Langer<sup>†</sup>

Sofi Sinozich<sup>‡</sup>

February 2018

*Forthcoming in Public Opinion Quarterly*

**Abstract:** This paper presents state-level estimates of the 2016 presidential election using data from the ABC News/Washington Post tracking poll and multilevel regression with poststratification (MRP). While previous implementations of MRP for election forecasting have relied on data from prior elections to establish poststratification targets for the composition of the electorate, in this paper we estimate both turnout and vote preference from the same pre-election poll. Through Bayesian estimation we are also able to capture uncertainty in both estimated turnout and vote preferences. This approach correctly predicts 50 of 51 contests, proving more accurate than comparison models that rely on the 2012 Current Population Survey Voting and Registration Supplement for turnout. While the model does not perfectly estimate turnout as a share of the voting age population, popular vote shares or vote margins in each state, it proves more accurate than predictions published by polling aggregators or other published MRP estimators. The paper also reports how vote preferences changed over the course of the 18-day tracking period, compares subgroup-level estimates of turnout and vote preferences with the 2016 CPS Survey and National Election Pool exit poll, and summarizes the accuracy of the approach applied to the 2000, 2004, 2008 and 2012 elections. The paper concludes by discussing how researchers can make use of this method an alternative approach to survey weighting and likely voter modeling as well as in forecasting future elections.<sup>§</sup>

---

<sup>\*</sup> Affiliated professor of the Division of Political Studies at the Centro de Investigación y Docencia Económicas (CIDE), Mexico City, Mexico, and a former senior research analyst at Langer Research Associates, New York, NY, USA. Please direct correspondence to [chadkdj@gmail.com](mailto:chadkdj@gmail.com).

<sup>†</sup> President of Langer Research Associates, New York, NY, USA.

<sup>‡</sup> Research analyst at Langer Research Associates, New York, NY, USA.

<sup>§</sup> Acknowledgements: The authors thank ABC News for access to the survey data and Robert Shapiro, Jon Krosnick, Andrew Gelman, the editor and three anonymous reviewers for valuable comments.

## **I. Overview**

Donald Trump's widely unexpected victory in the 2016 U.S. presidential election raised questions about the accuracy of public opinion polling, the aggregation of polling into probabilistic election forecasts and the interpretation of election polling by data analysts, journalists and the public. While national-level polls on average proved as accurate as in past elections in predicting the popular vote (with an average error on the margin of about 2 points), there were substantial state-level errors, particularly in Midwestern swing states (Enten 2016; Silver 2017; Cohn, Katz and Quealy 2016). These misses, amplified by a proliferation of overconfident forecasts, placed a cloud over the polling industry, provided ammunition to critics of survey research and led the leading industry association to study how the 2016 misses can be avoided in the future. To improve the accuracy of poll-based election forecasting, this paper proposes a new method for using multilevel regression with poststratification (MRP), a statistical technique developed over the past two decades and increasingly used for estimating public opinion, including electoral preferences, within subnational units from national-level survey data (Park, Gelman and Bafumi 2004; Lax and Phillips 2009a, 2009b; Pacheco 2011; Ghitza and Gelman 2013; Lauderdale 2016; Gelman and Little 1997).

In the electoral forecasting setting, MRP combines statistical modeling of national-level survey data with estimates of demographic group sizes at the state level to produce state-level candidate preference estimates. By combining individual-level demographic and state-level predictors, the technique leverages subgroup-level similarities in electoral attitudes across geographical units (in this case, states). Through this partial pooling approach, MRP can provide highly accurate state-level estimates of voter intentions even with relatively small state-level sample sizes.

One key challenge facing this approach is identifying accurate demographic targets on which to poststratify the model predictions. While using census data to estimate subnational public opinion among the general public is straightforward (e.g. Lax and Phillips 2009a, 2009b), the population of interest for election forecasts is voters, an unknown population. Previous studies have employed voter turnout estimates from the immediate prior election. For example, to estimate the 2012 election results, Wang et al. (2014) poststratified on turnout estimates from the 2008 National Election Pool (NEP) exit poll. Others have conducted MRP analyses using reported turnout data from the U.S. Census Bureau’s Current Population Survey (CPS) Voting and Registration Supplement (e.g. Lauderdale 2016). The disadvantage with this approach is that using prior turnout estimates assumes that changes in relative subgroup turnout between elections is fairly negligible, a risky assumption.

In this paper, we propose using pre-election survey data to estimate not only candidate preferences, but also turnout, using the latter to estimate poststratification targets for the former. This two-step method helps capture change in relative subgroup-level turnout between elections. Although this approach may be subject to bias due to overreported turnout intentions, as long as this bias is fairly constant across subgroups, the results should remain largely accurate (Ghitza and Gelman 2013). Further, by estimating both turnout and preference models using Bayesian estimation, we are able to propagate the uncertainty from the turnout estimates through to the final candidate preference estimates; most previous approaches have not taken into account uncertainty in turnout estimates.

We implement our MRP approach using survey data from the 2016 ABC News and ABC News/Washington Post tracking poll, a national random-digit dialed (RDD) survey of the general public. We developed the model experimentally during the election; minor subsequent

refinements included the use of more recent census data, elimination of a phone-status adjustment and addition of a time-trend adjustment. Although these changes had little effect, our results were not publicly released prior to the election. As such we present this analysis solely as a proof of the utility of the approach, not as a claim to have made a public prediction.

Using this dataset, our approach anticipated that Hillary Clinton would almost certainly win the popular vote (96 percent probability), but gave a slight edge to Trump in the Electoral College (64 percent probability). The models correctly predicted the presidential winner in 50 of the 51 states and the District of Columbia (missing only Wisconsin), estimated the national popular vote margin within half of a percentage point, and produced lower errors on the Clinton-Trump margin across states than leading polling aggregators and non-probability 50-state polls. By contrast, poststratifying on 2012 CPS estimates for turnout in the previous presidential election suggests large victories for Clinton in both the popular vote and Electoral College, reflecting differing turnout by racial and educational groups. Additionally, we obtain largely accurate results from analyses of ABC News and ABC/Post tracking data from the previous four presidential elections, demonstrating the robustness of the approach.

More than solely a predictive enterprise, our analysis also sheds light on the dynamics of the 2016 race, in terms of trends over time and voting behavior among demographic groups. Except for the first four days of the tracking poll, Oct. 20-23, the model found Trump ahead in the Electoral College even as he generally trailed Clinton in the popular vote. Notably, the narrowing of the race occurred prior to FBI Director James Comey's Oct. 28 letter to Congress reopening the controversy over Clinton's use of a private email server as secretary of state – an important analytical finding, given subsequent commentary blaming Clinton's loss on the Comey letter. We also find an older, whiter and less well-educated electorate than suggested by the NEP

2016 national exit poll, and a lower share of black voters than one might expect from 2012 CPS data, both key to understanding the 2016 electoral outcome.

## **II. Using MRP for State-level Turnout and Vote Estimates**

National-level polling plus MRP is a promising alternative for election forecasting, avoiding uncertain rigor in state polls and the need for prescience in anticipating where to conduct them. The statistical properties and substantive advantages of MRP have been discussed in detail elsewhere.<sup>1</sup> Broadly, researchers start with a national-level survey dataset, preferably with a substantial number of observations. A multilevel statistical model predicting the outcome of interest is fit, using basic demographic variables that are available in census data at the state (or other subnational unit) level. Additional state-level variables can be included in the model to sharpen estimates. Coefficients for continuous variables usually are unmodeled (i.e., fixed), while group variables are modeled as categorical random effects.

In the second stage of MRP analysis, the model estimates are used to predict the outcome variable for groups defined in a poststratification dataset. This dataset has an observation corresponding to each group defined for all combinations of the demographic variables included in the model. For example, if the model includes U.S. states/D.C. ( $n=51$ ), gender ( $n=2$ ) and a four-category race/ethnicity variable ( $n=4$ ), the poststratification dataset will have  $51*2*4=408$  rows, and will include the population size in each group, e.g., from census estimates. After predicting the outcome variable for each of the groups in the poststratification dataset, estimates

---

<sup>1</sup> See Park, Gelman and Bafumi (2004), Ghitza and Gelman (2013), Lax and Phillips (2009a, 2009b), Pacheco (2011), Warshaw and Rodden (2012), Buttice and Highton (2013), and Wang et al. (2014).

can be aggregated to the state (or other subgroup unit) level, with the subgroup population sizes determining the relative weight of each group's estimate in the state-level estimate.

MRP provides a powerful approach to generating state-level opinion estimates by pooling information from similar groups in other states, in effect leveraging subgroup-level attitudinal homogeneity. Through multilevel modeling, results among multiple groups are partially pooled; for example, the prediction for Hispanic men living in Oregon will be informed by the outcome among Hispanic men in other states, other people living in Oregon, men across the country, Hispanics across the country, etc.

The degree to which estimates are partially pooled across groups is largely dependent on sample sizes; with larger samples, group estimates more closely reflect the outcome in the data, while estimates for smaller sample size groups are more dependent on the model (i.e., information from similar groups). Thus, pooling of data across states is particularly important for states with smaller sample sizes; estimates from states with larger sample sizes rely more on the survey data and less on the statistical model.

As noted, we use two sets of MRP models to predict not only election preferences, but also turnout among demographic groups to establish poststratification targets for the former. Starting first with candidate preferences, two sets of multilevel logistic regression models predicting preference for Clinton or Trump, respectively, are fit, with respondents preferring the other major-party candidate, a third-party candidate, or undecided set to 0.

The candidate preference equations are as follows:

$$\begin{aligned}
\Pr(\text{candidate}_i^k) = & \text{logit}^{-1} \left( \alpha_0 + \beta_1(2012 \text{ party share})_j + \beta_2(\text{black share})_j + \right. \\
& \beta_3(\text{Hispanic share})_j + \beta_4(\text{white evang. share})_j + \alpha_{1[i]}^{\text{gender}} + \\
& \alpha_{2[i]}^{\text{age5}} + \alpha_{3[i]}^{\text{race4}} + \alpha_{4[i]}^{\text{edu5}} + \alpha_{5[i]}^{\text{state}} + \alpha_{6[i]}^{\text{region}} + \alpha_{7[i]}^{\text{age5,edu5}} + \\
& \alpha_{8[i]}^{\text{gender,edu5}} + \alpha_{9[i]}^{\text{race5,age5}} + \alpha_{10[i]}^{\text{race5,edu5}} + \alpha_{11[i]}^{\text{race5,gender}} + \\
& \left. \alpha_{12[i]}^{\text{race,region}} + \alpha_{13[i]}^{\text{wave}} \right) \tag{1}
\end{aligned}$$

In this equation,  $\alpha_0$  is the baseline intercept, while the remaining alpha coefficients correspond to variance components for the grouping variables, which include typical survey weighting variables (gender, age, education and race/ethnicity), geographical factors (state, region), and interactions (age\*education, gender\*education, race\*age, race\*education, race\*gender and race\*region). To capture any time trends, a random effect for time periods in the tracking survey is included. The equation also includes three state-level variables and associated fixed beta coefficients to increase the precision of state-level estimates suggested by previous research by others and testing with 2012 ABC/Post data. These include the Obama and Romney vote shares in each state, respectively, (Ghitza and Gelman 2013; Wang et al. 2014) and the proportion of each state's population made up of African-Americans, Hispanics and evangelical white Protestants.

Instead of using 2012 exit poll or CPS turnout estimates to poststratify the results of the preference models, we estimate turnout using expressed turnout intentions captured in the pre-election survey. Respondents who said they were registered to vote at their current address, would definitely vote and reported voting in 2012 or had voted early are classified as voters (=1) while all other respondents are non-voters (=0). Of three operationalizations of turnout that were tested, this approach produces the most accurate estimates using 2012 ABC/Post data. It should

be noted that the goal was not to assemble literal voters but rather to predict the probability of turnout among groups, something this simplified approach accomplishes effectively. As a result, turnout is estimated with a nearly identical multilevel logistic regression model, with the difference being that the only state level variable included is the voting age population turnout in each state in 2012:

$$\begin{aligned}
\Pr(\text{vote}_i) = \text{logit}^{-1} & \left( \alpha_0 + \beta_1 (2012 \text{ VAP turnout})_j + \alpha_{1[i]}^{\text{gender}} + \alpha_{2[i]}^{\text{age5}} + \alpha_{3[i]}^{\text{race4}} \right. \\
& + \alpha_{4[i]}^{\text{edu5}} + \alpha_{5[i]}^{\text{state}} + \alpha_{6[i]}^{\text{region}} + \alpha_{7[i]}^{\text{age5,edu5}} + \alpha_{8[i]}^{\text{gender,edu5}} + \alpha_{9[i]}^{\text{race5,age5}} \\
& \left. + \alpha_{10[i]}^{\text{race5,edu5}} + \alpha_{11[i]}^{\text{race5,gender}} + \alpha_{12[i]}^{\text{race,region}} + \alpha_{13[i]}^{\text{wave}} \right) \\
\alpha_{l[i]}^S & \sim N(0, (\sigma^S)^2)
\end{aligned} \tag{2}$$

In the next stage, the turnout model estimates were poststratified on the census dataset, producing an estimate of the number of Americans in each subgroup who were likely to vote. The preference model estimates (vote for Clinton, vote for Trump) next were poststratified on the estimated likely voter population in each subgroup. These estimates were adjusted to reflect time trends by using the final survey wave random effect estimate in the prediction, which essentially assumes uniform swings<sup>2</sup> in turnout/candidate preferences among groups. With these estimated numbers of voters overall and for each of the two main candidates, estimates of all three quantities could be aggregated to the state as well as subgroup levels.

---

<sup>2</sup> Consequential differential late-stage swings don't appear to be an issue, given the robustness of our model over time, and reliably testing time-trend interactions would require larger sample sizes than we have available.



### III. Data and Estimation

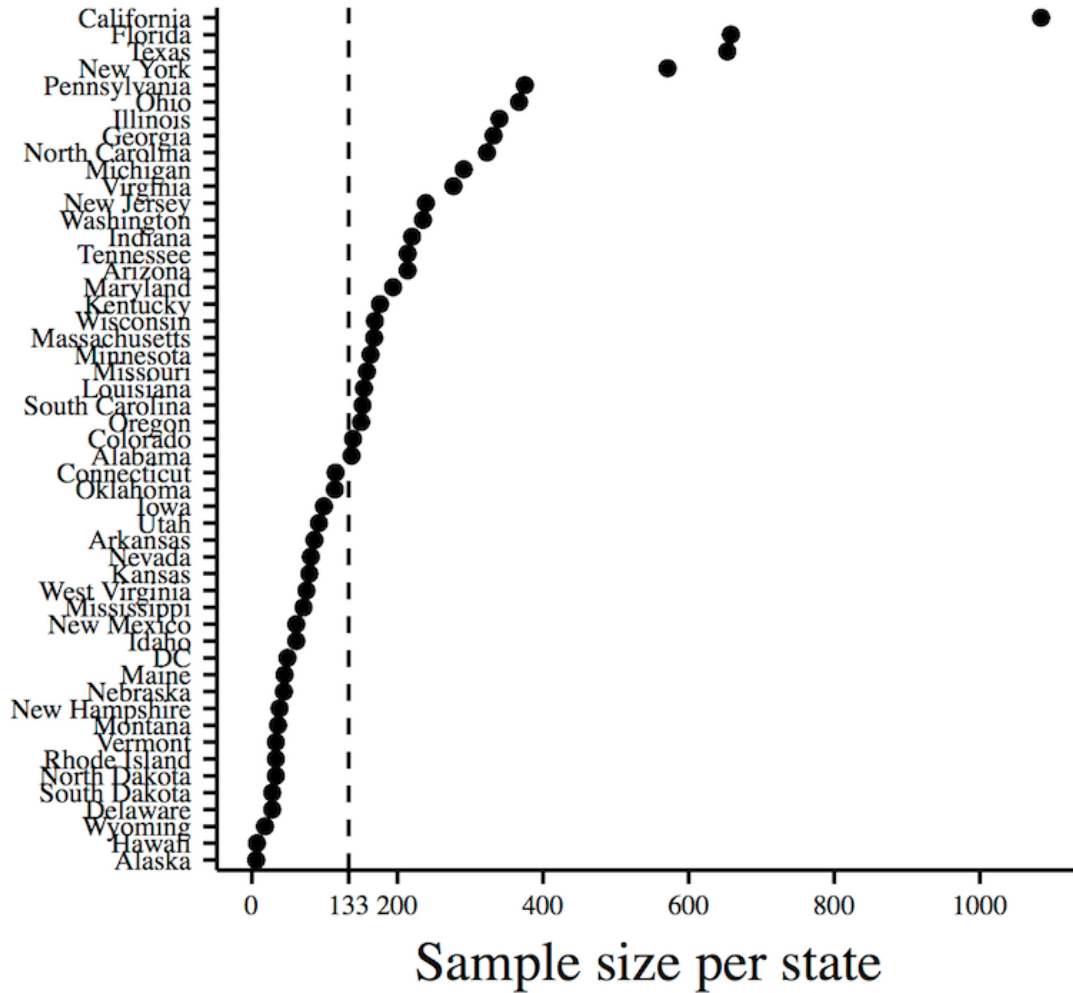
The national survey data chiefly used in the analysis are from the 2016 ABC News and ABC News/Washington Post tracking poll, a probability-based RDD cellular and landline telephone survey of respondents in the continental United States. The AAPOR response rate 3 for the full survey was 15.6 percent, including a cooperation rate (AAPOR 3) of 38.7 percent. We use data collected over 18 days preceding the Nov. 8, 2016, election, ending on Nov. 6. Approximately 440 members of the general public were interviewed during each of the first 14 days, and about 800 per day on days 15-18, Thursday-Sunday before the election. A total of 9,485 respondents (including 7,778 self-reported registered voters) are included in the dataset, 65 percent of whom were interviewed via cell phone.<sup>3</sup> Figure 1 plots the sample size by state, ranging from six respondents in Alaska<sup>4</sup> to 1,084 in California, with the median state (Colorado) including 133 respondents.

A total of 6,193 respondents were classified as likely voters using the definition described above. Candidate preference questions were asked only of respondents who said they had voted or definitely would vote ( $n=6,825$ ). Of these, those who said they preferred/voted for Clinton were coded as 1 for the Vote Clinton variable ( $n=3,073$ ), with all others (supporters of Trump, third-party candidates and undecideds) coded as 0. Similarly, for the Vote Trump variable, Trump early voters and supporters were coded 1 ( $n=3,005$ ) and all others 0.

---

<sup>3</sup> Additional methodological details are available at <http://abcnews.go.com/US/PollVault/abc-news-polling-methodology-standards/story?id=145373>.

<sup>4</sup> Respondents from Alaska and Hawaii only include those reached on cell phones with area codes from the continental United States.



**Figure 1.** General population sample sizes per state. The dashed line indicates the median state sample size. MRP turnout models used this full sample, while candidate preference models were estimated among those who said they had voted or definitely would vote.

Demographic group variables included as random effects were gender (male, female), age (18-29, 30-39, 40-49, 50-64, 65+), race/ethnicity (white non-Hispanic, black non-Hispanic, Hispanic, and other non-Hispanic), and education (less than high school, high school graduate, some college, four-year college graduate, post-graduate). State of residence was determined by area code and exchange for landline respondents and asked of cell phone respondents. The region variable used 13 sociopolitical regions based on state of residence. To capture any time

trends, the tracking period was divided into five periods, the first three of which were four days in length and the last two at three days each.

Poststratification data were taken from the U.S. Census Bureau's 2015 American Community Survey (ACS) one-year estimates, the most recent census data released before the election. Since the poststratification dataset includes cells for every combination of the demographic variables included in the model, the dataset contains 10,200 rows with ACS estimates of the population size in each of these groups.

State-level variables included past voting-age-population turnout estimates compiled by McDonald (2016), prior vote shares for Obama and Romney in 2012, aggregate racial group shares from the 2015 ACS estimates and estimates of evangelical white Protestants in each state from the Public Religion Research Institute's American Values Atlas (2016).

To assess the performance of our approach, we also estimated turnout using the 2012 CPS Voter Supplement Data, which includes 82,820 usable observations collected after the 2012 election. This dataset is considered highly accurate; though like most surveys it suffers from overreporting bias, such biases are likely to be small, and much less significant than biases in the NEP exit poll, which consistently shows much younger and more educated electorates compared with voter file data (Ghitza and Gelman 2013; McDonald 2007). We ran a similar MRP model to predict turnout using the 2012 CPS data and poststratified on the 2015 ACS to account for population change since 2012. The turnout variable was recoded and reweighted to account for overreporting and non-response bias according to the recommendations of Hur and Achen (2013) and as implemented by McDonald (2017), which align the CPS survey results with the actual results at the state level.

We estimated initial models with various demographic interactions using an approximate maximum likelihood estimator in the `glmer` function in the `lme4` package in R (Bates and Maechler 2009), with random effects with zero estimated variances removed from the models to aid convergence. After these initial model runs, final models were estimated through Bayesian techniques using Stan, through the `rstanarm` package in R (Stan Development Team 2016).<sup>5</sup> By using this Bayesian approach, it is possible to account for uncertainty in both estimated turnout and vote preferences.<sup>6</sup>

#### **IV. 2016 MRP election estimates**

##### *Estimating turnout from pre-election polls outperforms models based on historical data*

To assess our method for estimating turnout, we compare state-by-state turnout predictions from the MRP model using the 2016 ABC/Post data with the 2012 CPS-based MRP model. Figure 2 plots the predicted turnout by state (y-axis) vs. the actual turnout (x-axis); since the CPS filters on citizenship status, “actual turnout” for the CPS model is based on turnout among the voting eligible population (VEP), while for the ABC/Post predictions turnout is plotted for the voting age population (VAP) (McDonald 2016).<sup>7</sup> Compared with actual turnout,

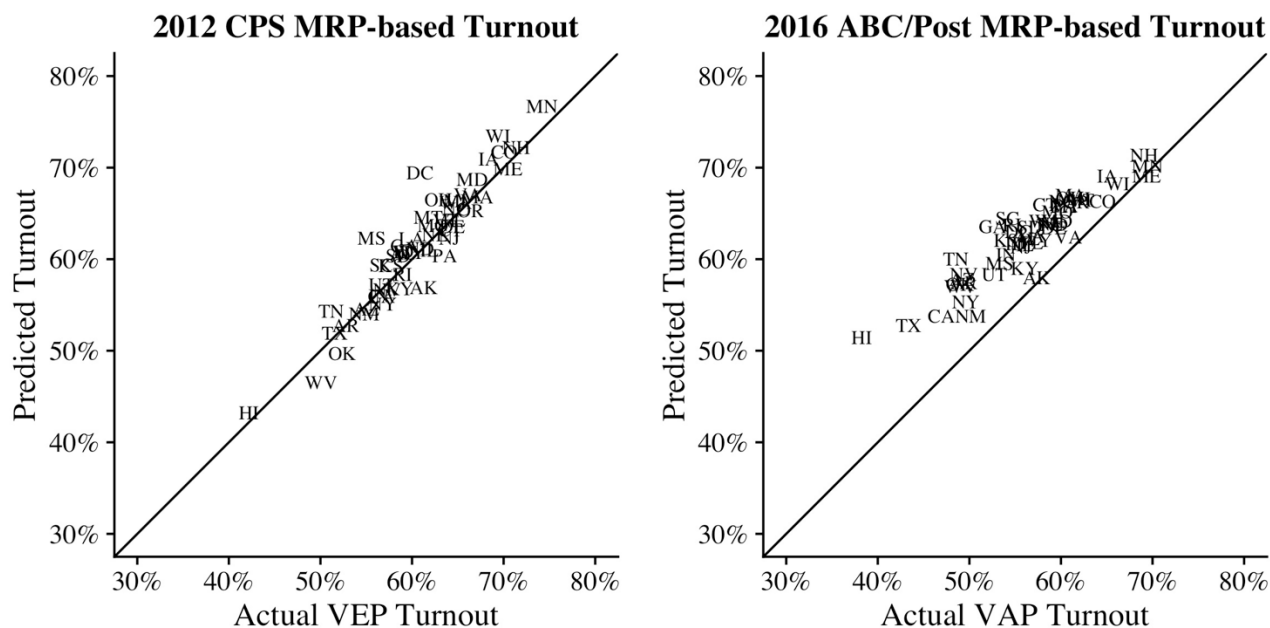
---

<sup>5</sup> Due to available computing power combined with the size of the CPS dataset, CPS turnout models were only estimated using the maximum likelihood estimator. This likely had little impact on the model point estimates, but means that final model predictions using CPS data understate uncertainty.

<sup>6</sup> The Bayesian models were estimated using `rstanarm`’s default weakly informative priors. Priors for the intercept and fixed coefficients are normally distributed with location parameters set to zero and scale parameters to 10 and 2.5, respectively. The prior on the covariance matrix sets the regularization, concentration, shape and scale parameters to 1. See [mc-stan.org/rstanarm/reference/priors.html](http://mc-stan.org/rstanarm/reference/priors.html).

<sup>7</sup> A more detailed discussion of the turnout results, including a state-by-state table of estimates is available in the online Appendix.

the 2012 CPS-based state-level predictions are generally close to the 45-degree line. This is not surprising, since the CPS data are weighted to the actual 2012 results by state. By contrast, the ABC/Post-based model generally overpredicts turnout, with absolute errors reaching double digits in a few states, including Hawaii (13 pts.), Georgia (11), Tennessee (11) and South Carolina (10). Turnout errors decrease as the level of actual turnout in each state increases; errors are minimal in high-turnout states such as Minnesota and Wisconsin. Indeed, the prediction error is correlated at  $r = -0.76$  with actual turnout. This outcome may reflect a greater likelihood of voters to answer surveys and a lack of overreporting adjustment available for the ABC/Post data.



**Figure 2.** Predicted state-by-state turnout in the 2016 election vs. actual turnout. The left panel presents predictions based on an MRP model of 2012 CPS data of citizens plotted against the actual VEP turnout. The right panel present predictions of an MRP model that utilizes the 2016 ABC/Post tracking survey to predict turnout among the general population and is plotted against the actual VAP turnout.

What is most important, however, is the relative accuracy of the predictions; from this perspective both the ABC/Post- and CPS-based model predictions are highly accurate; they correlate with actual turnout by state at  $r=0.92$  and  $r=0.94$ , respectively. Furthermore, turnout prediction errors for the ABC/Post-based models are essentially uncorrelated with the actual

state-by-state Clinton-Trump margins ( $r=0.004$ ); the same correlation for the CPS-based model errors is somewhat higher at  $r=0.19$ .

Next, we show the estimated turnout shares among key demographic subgroups predicted by both modeling approaches and compare them with the 2016 CPS and exit poll results. We expect that the 2016 CPS results represent the most accurate available estimate of the electorate, given known demographic biases in the NEP exit polls (Cohn 2016; McDonald 2007; Ghitza and Gelman 2013).

As shown in Table 1, turnout share estimates from the ABC/Post MRP are fairly similar to those produced by the 2012 CPS MRP model and in the 2016 CPS results, with a few important exceptions. First, the ABC/Post-based turnout model anticipates that blacks would make up only 11.3 percent of the electorate, similar to the 2016 CPS estimates (11.8 percent), but 2.7 points lower than expected by the 2012 CPS MRP results. Instead, both the ABC/Post turnout estimates and the 2016 CPS suggest more whites in the electorate, 73.0 and 73.6 percent, vs. 72.1 percent from the 2012 CPS MRP (and 71 percent in the exit poll). Second, relative to the 2016 CPS estimates, both MRP approaches overstate the share of voters without a college degree by 2 to 3 points; however, they both are much closer to the 2016 CPS estimates than the exit poll, which suggests that about half of the electorate had college degrees. Third, the ABC/Post MRP model predicts a somewhat smaller share of 18- to 29-year-olds than the 2012 CPS MRP, the 2016 CPS topline and the 2016 exit polls.<sup>8</sup> Given the importance of race and education in explaining vote preferences, combined with the more minor role of age (see below), the differing

---

<sup>8</sup> Though the 2012 exit poll-based estimates are similar to ABC/Post MRP in terms of gender and race, the exit poll estimates suggest a much younger and more educated electorate.

electorates produced by the two pre-election approaches have important ramifications for the overall model predictions.

|                 | <b>Pre-election<br/>Poststratification Targets</b> |                         | <b>Post-election Estimates</b> |                                |
|-----------------|--|-------------------------|--------------------------------|--------------------------------|
| <b>Subgroup</b> | <b>2016<br/>ABC/Post<br/>MRP</b>                   | <b>2012 CPS<br/>MRP</b> | <b>2016 CPS</b>                | <b>2016 NEP<br/>Exit Polls</b> |
| Male            | 46.6%  | 46.2%                   | 46.4%                          | 47.5%                          |
| Female          | 53.4   | 53.8                    | 53.6                           | 52.5                           |
| 18-29           | 11.1   | 14.9                    | 15.7                           | 19.1                           |
| 30-39           | 15.5   | 14.5                    | 15.0                           | 16.6                           |
| 40-49           | 18.0   | 16.5                    | 16.2                           | 18.7                           |
| 50-64           | 30.8   | 29.9                    | 28.8                           | 30.0                           |
| 65+             | 24.6   | 24.2                    | 24.2                           | 15.6                           |
| No degree NET   | 63.4   | 62.5                    | 60.4                           | 49.9                           |
| HS or less      | 29.7   | 29.0                    | 29.7                           | 18.3                           |
| Some college    | 33.7   | 33.6                    | 30.7                           | 31.6                           |
| Degree NET      | 36.6   | 37.5                    | 39.6                           | 50.1                           |
| College         | 22.9   | 23.1                    | 25.0                           | 31.9                           |
| Postgraduate    | 13.8   | 14.4                    | 14.6                           | 18.2                           |
| Whites          | 73.0   | 72.1                    | 73.6                           | 70.5                           |
| Nonwhites NET   | 27.0   | 27.9                    | 26.4                           | 29.5                           |
| Blacks          | 11.3   | 14.0                    | 11.8                           | 11.9                           |
| Hispanics       | 9.1  | 8.7                     | 9.1                            | 11.2                           |
| Other/mixed     | 6.6  | 5.2                     | 5.5                            | 6.4                            |
| Among whites    |  |                         |                                |                                |
| Men no deg.     | 21.1   | 19.7                    | 20.0                           | 16.4                           |
| Women no deg.   | 23.8   | 22.7                    | 22.3                           | 17.0                           |
| Men deg.        | 13.7   | 14.3                    | 14.9                           | 17.3                           |
| Women deg.      | 14.4   | 15.4                    | 16.4                           | 19.9                           |

**Table 1.** Predicted turnout shares by demographic groups. The first two columns are from MRP model estimates (medians) using the 2016 ABC/Post data and 2012 CPS data, respectively. The remaining columns report topline estimates of turnout shares from the 2016 CPS and the 2016 NEP Exit Poll.

*MRP based on pre-election polling anticipates Trump victory, 2012 turnout based models don't*

To generate state-by-state vote forecasts, we next poststratified models predicting preferences for Clinton and Trump on the two electorates produced by the two MRP turnout models. Before even examining the state-by-state results, it is clear from the national level estimates that using the pre-election ABC/Post data to estimate the composition of the electorate proves both quite accurate and performs better than the CPS-based approach. Clinton won the national popular vote by 2.1 points (48.2 to 46.1 percent); the ABC/Post approach estimates a 2.6-point Clinton victory in the popular vote (46.8 to 44.3 percent), vs. a 5.8-point Clinton win using the 2012 CPS.

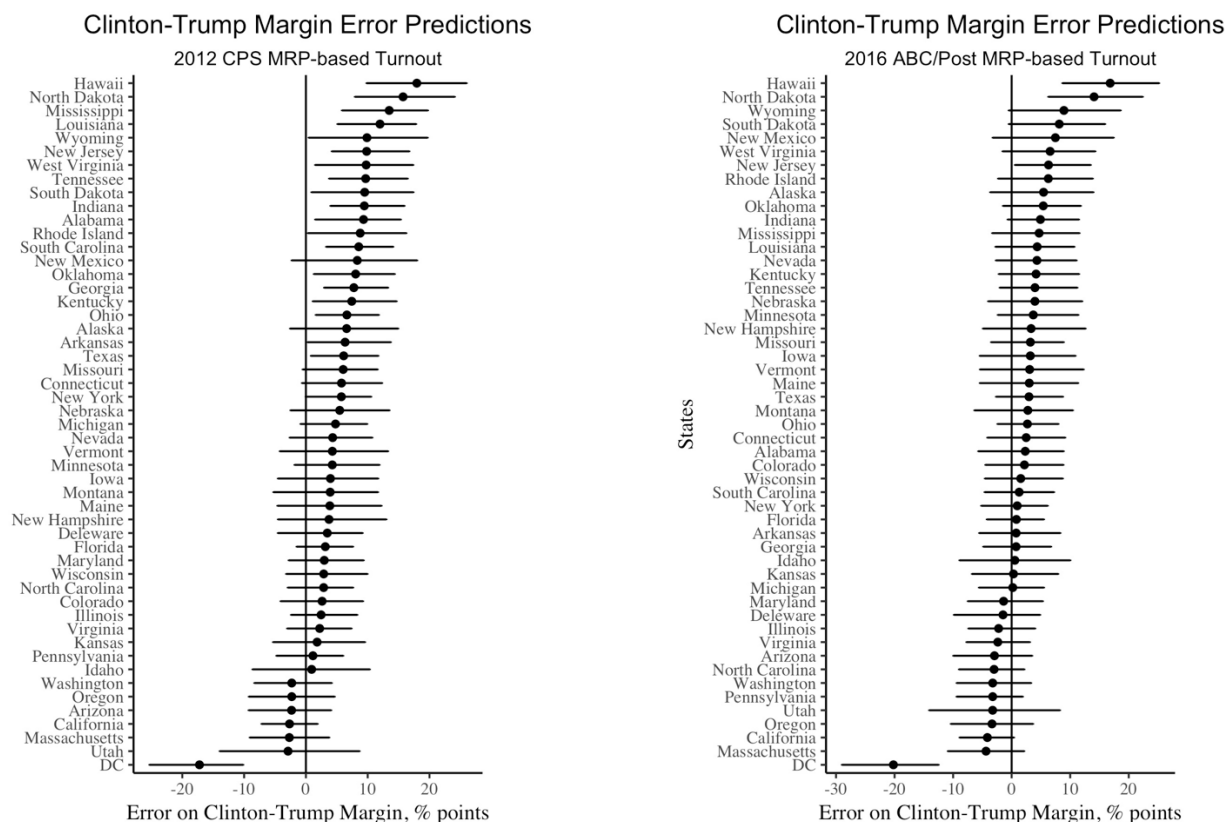
Although both data sources suggest a Clinton victory in the popular vote, the ABC/Post survey-based approach does so more accurately, and it alone correctly anticipates a Trump victory in the Electoral College. Figure 3 plots the state-by-state model estimates of the Clinton-Trump margin relative to the actual margin (i.e., the prediction errors).<sup>9</sup> The ABC/Post-based turnout proved highly accurate, correctly identifying the winner in 50 of 51 contests, missing just in Wisconsin, with only four of the predicted margins outside of the 95 percent credible intervals. For all states, the root mean square error (RMSE) on the Clinton-Trump margin is 5.6 points, dropping to 4.3 points after excluding Hawaii, Alaska and the District of Columbia. The absolute errors on the Clinton-Trump margins reach double digits only in two low-population states (Hawaii and North Dakota) and D.C., while eight state margin estimates are off by 5 to 10 points (Alaska, New Jersey, New Mexico, Oklahoma, Rhode Island, South Dakota, West

---

<sup>9</sup> Tables of model estimates are available for reference in the Appendix, along with additional models that vary the inclusion of state-level variables.



Virginia, and Wyoming). Furthermore, the RMSE for the Clinton and Trump estimates per state for the full model are 2.2 and 3.5 points (excluding Alaska, Hawaii and D.C.), respectively, indicating quite accurate point estimates for each candidate. Coverage rates for the 95 percent intervals on the Clinton and Trump vote share estimates are 90.1 percent for each, indicating a generally well-calibrated model, though with some slight underestimation of uncertainty.



**Figure 3.** Each panel plots the model errors on the Clinton-Trump margin by state, with perfect predictions at zero, overprediction for Clinton to the right of zero and overprediction for Trump to the left of zero. While preferences in both panels are estimated from the 2016 ABC/Post tracking survey data, the left panel estimates utilized an MRP model of the 2012 CPS Voter Supplement Survey to establish poststratification targets, while the right panel estimates used poststratification targets from an MRP turnout model estimated on the 2016 ABC/Post tracking data.

By contrast, using the 2012 CPS to estimate poststratification targets leads to incorrect predictions of Clinton victories in Florida, Georgia, Michigan, Pennsylvania and Wisconsin and produces larger RMSEs. The RMSE for all 51 contests reaches 7.4 points, or 6.7 excluding

Alaska, Hawaii, and D.C. While both approaches are essentially equally accurate in terms of the Clinton point estimates, the RMSE on the Trump estimates reaches 5.0 points for the CPS-based models, vs. 3.5 points for the ABC/Post based-turnout.

|              |               | <b>Huffpost<br/>Pollster</b> |              | <b>ABC/Post-based turnout MRP<br/>Estimates</b> |              |                              | <b>2012 CPS-based turnout MRP<br/>Estimates</b> |              |                              |
|--------------|---------------|------------------------------|--------------|---|--------------|------------------------------|---|--------------|------------------------------|
| <b>State</b> | <b>Actual</b> | <b>Average</b>               | <b>Error</b> | <b>Predicted</b>                                | <b>Error</b> | <b>Clinton<br/>win prob.</b> | <b>Predicted</b>                                | <b>Error</b> | <b>Clinton win<br/>prob.</b> |
| Michigan     | -0.2          | 6.0                          | 6.2          | 0.0   | 0.2          | 49%                          | 4.6   | 4.8          | 95%                          |
| Georgia      | -5.2          | -2.4                         | 2.8          | -4.4  | 0.8          | 7%                           | 2.6   | 7.8          | 85%                          |
| Florida      | -1.2          | 1.8                          | 3.0          | -0.4  | 0.8          | 44%                          | 1.9   | 3.1          | 80%                          |
| Wisconsin    | -0.8          | 6.1                          | 6.9          | 0.8   | 1.6          | 59%                          | 2.1   | 2.9          | 75%                          |
| Colorado     | 4.9           | 4.9                          | 0.0          | 7.1   | 2.2          | 98%                          | 7.5   | 2.6          | 98%                          |
| Virginia     | 5.3           | 5.3                          | 0.0          | 3.0   | -2.4         | 86%                          | 7.6   | 2.2          | 100%                         |
| Ohio         | -8.1          | -1.0                         | 7.1          | -5.4  | 2.7          | 2%                           | -1.5  | 6.6          | 28%                          |
| Arizona      | -3.5          | -1.6                         | 1.9          | -6.5  | -2.9         | 2%                           | -5.9  | -2.3         | 4%                           |
| N. Carolina  | -3.7          | 1.6                          | 5.3          | -6.7  | -3.0         | 1%                           | -0.8  | 2.9          | 37%                          |
| Penn.        | -0.7          | 4.1                          | 4.8          | -4.0  | -3.2         | 7%                           | 0.4   | 1.1          | 56%                          |
| Iowa         | -9.4          | -3.0                         | 6.4          | -6.2  | 3.2          | 5%                           | -5.4  | 4.0          | 8%                           |
| N. Hamp.     | 0.4           | 3.3                          | 2.9          | 3.7   | 3.3          | 81%                          | 4.1   | 3.7          | 83%                          |
| Minnesota    | 1.5           | 6.9                          | 5.4          | 5.2   | 3.7          | 95%                          | 5.8   | 4.3          | 97%                          |
| Nevada       | 2.4           | 2.1                          | -0.3         | 6.7   | 4.3          | 97%                          | 6.8   | 4.3          | 97%                          |

**Table 2.** Actual Clinton-Trump margins vs. the Huffpollster polling average and MRP estimates in swing states. Model win probabilities for the predictions based on the 2012 CPS turnout model do not account for uncertainty in turnout.

Notably, the ABC/Post turnout based MRP estimates are particularly accurate in swing states, with the RMSE on the Clinton-Trump margin only 2.5 points (vs. 4.1 points for the CPS-based estimates), proving more accurate in the states that mattered most in deciding the election outcome. Table 2 reports the predicted and actual margins in swing states along with margin errors and polling averages from Huffpost Pollster for comparison. The ABC/Post MRP estimate came within 1 point of the actual Clinton-Trump margin in Michigan, Georgia, and Florida; 1 to 3 points in Wisconsin, Colorado, Virginia, Ohio, and Arizona; and 3 to 4.3 points in North Carolina, Pennsylvania, Iowa, New Hampshire, Minnesota, and Nevada. While the polling

averages were particularly poor in Midwestern swing states with high numbers of less-educated whites, in Colorado, Virginia, Arizona, New Hampshire and Nevada the polling average is closer than the ABC/Post-based MRP estimates.

*Model estimates suggest an electorate even more polarized by education than the exit poll*

The models predicting candidate preferences clearly indicate group differences (see Table B1 in the Appendix). Variance in support for both Clinton and Trump is best explained by race/ethnicity, followed by gender, race by education, race by gender and education on its own. Age is somewhat less important; it does not explain much of the support for Clinton and has less impact than other variables on support for Trump. Higher support for Clinton and lower support for Trump among younger generations is more a function of their higher levels of education and greater nonwhite shares.

In terms of the fixed effects, Obama and Romney vote shares in the respondents' states are highly predictive, as is the proportion of each state's residents who are black, with higher shares related to lower support for Clinton and higher support for Trump, reflecting the greater GOP lean of whites in states with higher percentages of black residents. There's a similar pattern for the Hispanic share of residents, though this variable only reaches marginal levels of significance in the Trump support model. The influence of state-level proportions of evangelical white Protestants does not reach statistical significance in either model.

These patterns emerge in subgroup-level candidate preference predictions and underline the pitfalls of using turnout in a prior election to establish postratification targets. Table 3 compares estimates for subgroup-level candidate preferences from the two MRP estimates (2016 ABC/Post and 2012 CPS turnout) with estimates from the national exit poll. Given that the only difference between the two MRP models is the source of turnout poststratification targets, it is

|               | Clinton                 |                    |                                      | Trump                   |                    |                                      | Margin                  |                    |                                      |
|---------------|-------------------------|--------------------|--------------------------------------|-------------------------|--------------------|--------------------------------------|-------------------------|--------------------|--------------------------------------|
| Subgroup      | 2016<br>ABC/Post<br>MRP | 2012<br>CPS<br>MRP | 2016<br>NEP Exit<br>Polls<br>Topline | 2016<br>ABC/Post<br>MRP | 2012<br>CPS<br>MRP | 2016<br>NEP Exit<br>Polls<br>Topline | 2016<br>ABC/Post<br>MRP | 2012<br>CPS<br>MRP | 2016<br>NEP Exit<br>Polls<br>Topline |
| Male          | 40.4                    | 41.8               | 40.9                                 | 50.0                    | 48.3               | 51.7                                 | -9.6                    | -6.5               | -10.8                                |
| Female        | 52.4                    | 53.9               | 54.0                                 | 39.2                    | 37.6               | 40.9                                 | 13.2                    | 16.3               | 13.1                                 |
| 18-29         | 50.0                    | 52.7               | 54.6                                 | 33.3                    | 31.0               | 35.4                                 | 16.7                    | 21.7               | 19.2                                 |
| 30-39         | 50.9                    | 52.9               | 51.3                                 | 36.0                    | 34.2               | 39.2                                 | 14.9                    | 18.7               | 12.1                                 |
| 40-49         | 47.7                    | 49.1               | 46.1                                 | 44.4                    | 43.0               | 48.3                                 | 3.3                     | 6.1                | -2.2                                 |
| 50-64         | 45.7                    | 46.7               | 44.0                                 | 46.8                    | 45.7               | 52.2                                 | -1.1                    | 0.9                | -8.2                                 |
| 65+           | 43.5                    | 44.3               | 44.9                                 | 51.2                    | 50.4               | 51.7                                 | -7.7                    | -6.2               | -6.8                                 |
| No degree NET | 42.6                    | 44.8               | 43.6                                 | 49.0                    | 46.5               | 50.7                                 | -6.3                    | -1.7               | -7.1                                 |
| HS or less    | 42.8                    | 45.7               | 45.5                                 | 50.0                    | 46.8               | 50.4                                 | -7.2                    | -1.1               | -4.9                                 |
| Some college  | 42.5                    | 44.0               | 42.5                                 | 48.1                    | 46.2               | 50.8                                 | -5.6                    | -2.3               | -8.3                                 |
| Degree NET    | 54.1                    | 54.2               | 51.9                                 | 36.1                    | 36.0               | 41.4                                 | 17.9                    | 18.2               | 10.5                                 |
| College       | 50.1                    | 50.1               | 48.6                                 | 39.4                    | 39.2               | 44.2                                 | 10.7                    | 10.9               | 4.4                                  |
| Postgraduate  | 60.7                    | 60.8               | 57.8                                 | 30.6                    | 30.7               | 36.6                                 | 30.0                    | 30.1               | 21.2                                 |
| Whites        | 37.7                    | 38.4               | 37.2                                 | 53.6                    | 52.6               | 56.4                                 | -15.9                   | -14.2              | -19.2                                |
| Nonwhites NET | 71.4                    | 73.9               | 72.0                                 | 19.0                    | 16.6               | 22.3                                 | 52.5                    | 57.3               | 49.7                                 |
| Blacks        | 86.4                    | 86.7               | 88.6                                 | 5.7                     | 5.3                | 8.0                                  | 80.7                    | 81.4               | 80.6                                 |
| Hispanics     | 63.0                    | 63.4               | 65.6                                 | 27.6                    | 26.8               | 28.0                                 | 35.4                    | 36.7               | 37.6                                 |
| Other/mixed   | 57.5                    | 57.1               | 61.2                                 | 29.7                    | 29.6               | 30.9                                 | 27.8                    | 27.5               | 30.3                                 |
| Among whites  |                         |                    |                                      |                         |                    |                                      |                         |                    |                                      |
| Men no deg.   | 25.5                    | 25.7               | 22.7                                 | 65.9                    | 65.4               | 70.9                                 | -40.4                   | -39.7              | -48.2                                |
| Women no deg. | 34.9                    | 35.1               | 34.0                                 | 57.6                    | 57.1               | 60.9                                 | -22.7                   | -22.0              | -26.9                                |
| Men deg.      | 43.6                    | 43.7               | 38.4                                 | 45.8                    | 45.6               | 53.0                                 | -2.2                    | -1.9               | -14.6                                |
| Women deg.    | 54.6                    | 54.8               | 50.9                                 | 36.4                    | 36.1               | 43.4                                 | 18.3                    | 18.7               | 7.5                                  |

**Table 3.** MRP subgroup estimates of candidate preferences compared with NEP national exit poll estimate

unsurprising that the candidate preference predictions are highly similar. Predicted Clinton-Trump margins differ noticeably only where differences in turnout among blacks (and whites) are most relevant: among those with lower levels of education and younger voters. In both cases, the 2012 CPS-based MRP models predict more Clinton voters and fewer Trump voters.

Estimates from both MRP variants differ substantially with those from the NEP exit poll. We find greater polarization in the vote by education, overall and among whites only. Our ABC-Post MRP analysis finds that Clinton won college graduates by 18 points, vs. 10 points in the exit poll. Among whites, Trump won men and women without a college degree by slightly narrower margins according to our preferred MRP relative to the exit poll, while Clinton was far closer among white men with a college degree (-2 vs. -14 Clinton-Trump) and was farther ahead among white women with a college degree (+18 vs. +7).

Put together, the exit poll suggests that whites were more pro-Trump than the MRP models find, going for the Republican candidate by 20 points, vs. 16 points according to the ABC/Post MRP. Clinton's margin of victory among blacks and Hispanics is similar in both approaches. Notably, our preferred MRP model suggests that Clinton won Hispanics by 35 points, close to the exit poll's 38 points but lower than some have argued using other data sources (Sanchez and Barreto 2016).

The MRP estimates may be more accurate than the exit poll, which resorts to observation-based adjustments to differential nonresponse by gender, race and age; weights otherwise simply are used to align the data with actual vote results. Previous analyses have suggested that these non-response adjustments are insufficient to compensate for higher exit poll cooperation rates among more educated and younger voters (McDonald 2007; Cohn 2016). This can lead to implausible estimates for turnout, particularly half of the electorate having college

degrees vs. only about three in 10 Americans overall. By forcing the data to “add up” to the actual election results, the exit poll likely underestimated Clinton’s support among whites with college degrees and slightly overestimated Trump’s support among less-educated whites.

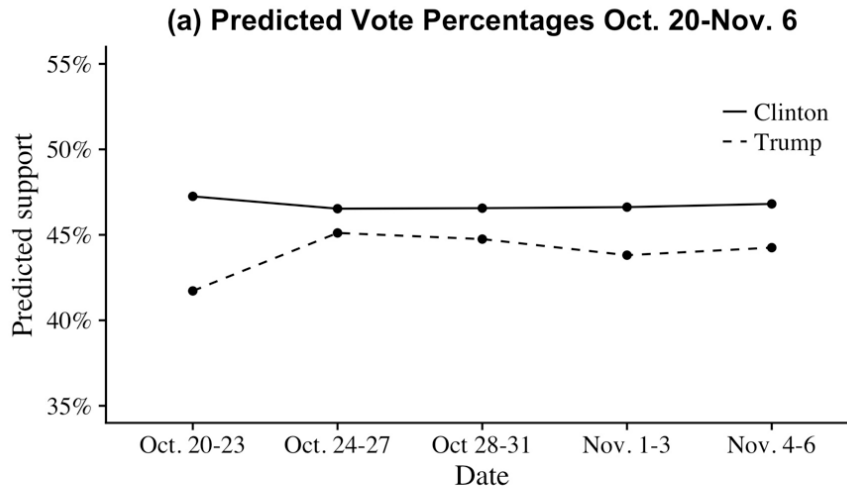
*Clinton consistently led in the popular vote, but not in the electoral vote*

Our ABC/Post MRP model predictions suggest that aside from a larger Clinton lead in the popular vote in the first four days of the tracking period (Oct. 20-23), the race was quite close, with Clinton ahead in the popular vote but Trump leading in the Electoral College. Figure 4 plots estimated Clinton-Trump popular and Electoral College vote shares using the full tracking data adjusted for the random effect predictions for each three- to four-day period.<sup>14</sup>

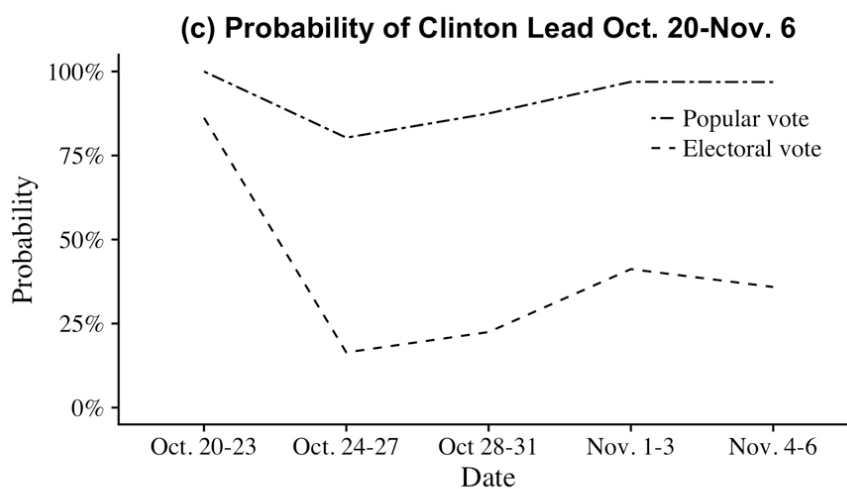
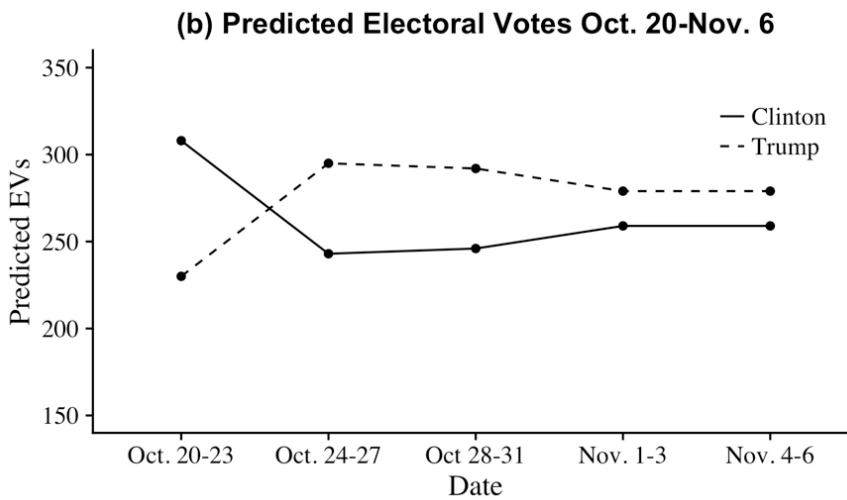
This approach suggests that Clinton was leading by 5.7 points in the popular vote during the first four days of the tracking period and by 308-230 Electoral College votes. The race then narrows substantially; though Clinton retained the support of 46.5-46.8 percent of voters throughout the full course of tracking, with the model predicting with 80 percent probability or more that she was leading in the popular vote, Trump came as close as 45.1 percent in the second period (Oct. 24-27). However, Trump led Clinton in predicted Electoral College votes during every period beyond the first four days, with Electoral College win probabilities ranging from 56 percent Nov. 1-3 to 83 percent Oct. 24-27. Notably, the narrowing of the contest at the national level, and the flip in the Electoral College prediction, occurred prior to FBI Director James Comey’s Oct. 28 letter to Congress about a renewal of the agency’s investigation into Clinton’s email server.

---

<sup>14</sup> Estimating models for each wave separately results in more dramatic period-to-period shifts, though the overall story remains similar.



**Figure 4.** MRP candidate support predictions across the tracking period, with (a) popular vote predictions (b) electoral vote predictions and (c) Clinton lead probabilities.



## **V. MRP proves more accurate than polling aggregators**

While ABC/Post-based MRP accuracy is impressive, given the closeness of the election in many states and the RMSEs of the estimates, missing only one state likely was due to chance. However, as shown in Table 4, our ABC/Post-based MRP model results outperform the predictions of polling aggregators as well as YouGov’s MRP model that employed a large non-probability online dataset for vote preferences and the 2012 CPS for turnout. This is true not only in terms of the number of states correctly predicted but also in the accuracy of state-level point estimates.

As noted, our preferred MRP model correctly predicts 50 of 51 contests, vs. 46 correct predictions for the leading aggregators and 43 for YouGov’s MRP model. Similarly, the RMSE on the Clinton-Trump margin for our MRP model is 5.6 points for all states, dropping to 4.3 points after excluding Alaska, Hawaii and D.C. and 2.7 points among battlegrounds. In the comparison models, the RMSE exceeds 7 points for all states, with no or only marginal improvement when excluding Alaska, Hawaii and D.C.<sup>15</sup> Our preferred MRP model’s RMSE among battleground states also outperforms the comparison models, with the closest being FiveThirtyEight, with a 4.0-point RMSE for these states (vs. our 2.7). This greater accuracy also holds when examining the two-party margin RMSE as well as the estimates for each major-party candidate.

---

<sup>15</sup> The improvements in our MRP estimates for these contests reflect the fact that Alaska and Hawaii were not in the sample frame, while D.C. is an unusual case with few respondents.



|                                | ABC/Post<br>MRP | 538      | HuffPo   | DKos | Survey<br>Monkey | NY<br>Times<br>Upshot | YouGov<br>MRP |
|--------------------------------|-----------------|----------|----------|------|------------------|-----------------------|---------------|
| Clinton %                      | 46.8%           | 48.5%    | 45.7%    | NA   | 47.0%            | NA                    | 47.9%         |
| Trump %                        | 44.3%           | 44.9%    | 40.8%    | NA   | 41.0%            | NA                    | 44.1%         |
| Margin                         | 2.6 pts.        | 3.6 pts. | 4.9 pts. | NA   | 6.0 pts.         | NA                    | 3.8 pts.      |
| Clinton EVs                    | 259             | 302      | 323      | 323  | 334              | 323                   | 317           |
| Trump EVs                      | 279             | 235      | 215      | 215  | 188              | 197                   | 221           |
| Clinton win probability        | 36%             | 71%      | 98%      | 88%  | NA               | 85%                   | NA            |
| Correct state predictions      | 50              | 46       | 46       | 46   | 44               | 45*                   | 43*           |
| RMSE margin all                | 5.6%            | 7.0%     | 7.2%     | 7.0% | 7.4%             | 7.0%                  | 7.6%          |
| RMSE margin no AK, HI, DC      | 4.3%            | 6.7%     | 7.2%     | 7.0% | 7.4%             | 7.0%                  | 7.6%          |
| RMSE margin battlegrounds      | 2.7%            | 4.0%     | 4.6%     | 4.8% | 5.1%             | 4.9%                  | 5.6%          |
| RMSE Clinton % (no AK, HI, DC) | 2.2%            | 3.3%     | 3.6%     | 2.8% | 3.1%             | NA                    | 3.4%          |
| RMSE Trump % (no AK, HI, DC)   | 3.5%            | 4.1%     | 6.8%     | 6.5% | 6.6%             | NA                    | 4.7%          |

**Table 4.** ABC/Post based MRP prediction estimates and RMSEs compared with others. Note that the predicted Electoral Votes for each candidate for our ABC/Post MRP model, the FiveThirtyEight model, and the YouGov MRP model do not necessarily correspond to the state-by-state popular vote estimates as they represent medians across model simulations, which do not always match the popular vote winners in each state. \*States that are listed as ties are counted as missed predictions.

In particular, our model’s strength comes more from its substantially higher accuracy in estimating Trump’s vote share. The RMSE for our Clinton estimates is 2.2 points, compared with 2.8 to 3.6 points for the comparison estimates. The difference is larger on average for the Trump estimates, with the RMSE for MRP at 3.5 points, vs. comparison model RMSEs ranging from 4.1 to 6.8 points.

## VI. MRP performs fairly well in past elections

As quasi-out-of-sample tests, we conducted similar analyses using the ABC/Post tracking surveys for the 2000, 2004, and 2008 elections.<sup>16</sup> (We used the 2012 survey data to help build the 2016 models.) We term these as “quasi” since it is hard to know how analysts would have built the models prior to each election. On one hand, we implemented these models post-hoc, knowing the outcomes. To guard against fitting to the election results, we took a highly similar approach as we did for 2016 (and 2012) with regard to aggregate-level variables, with the exception that the aggregate-level white evangelical Protestant variable was not available for 2000 and 2004. (We instead include a variable for the aggregate proportion of “other” racial/ethnic groups in each state.) On the other hand, hypothetical researchers likely would have included different variables to account for known political dynamics during these elections; for example, in prior elections, income, rather than education, better predicted electoral choices (Gelman 2008). To partially account for this, the demographic interactions included in the final models also differed somewhat from the 2016 models; as with the 2012 and 2016 models, we tested a number of demographic interactions in initial models and included those with non-zero random effects estimates in the final models. Importantly, the inclusion or exclusion of such interactions was based on their predictive power in the models (i.e., predicting *individual* turnout and vote preferences), not on the resulting state-level poststratified predictions.

Table 5 summarizes the results for the last five presidential elections. Our MRP approach correctly estimates an essentially tied race in 2000, but reverses the popular vote and Electoral

---

<sup>16</sup> The 2000 and 2004 tracking polls included more sample days and more respondents per day than the latter tracking polls. For a clearer comparison, for these two elections a more limited number of tracking days were analyzed.

College winners. The models slightly overestimate Obama in the popular vote and Electoral College in 2008, while underestimating his Electoral College victory in 2012. The 2004 MRP slightly underestimates Bush's popular vote and Electoral College victories.

The models are not quite as accurate for 2000-2008 as for 2016 and 2012 (the latter by design). However, the models miss the Electoral College and popular vote winner only in the 2000 election. The average RMSE on the Democratic-Republican candidate margins across the elections excluding 2012 in all 51 contests is 7.0 points, falling to 6.0 points when Alaska, Hawaii and D.C. are excluded. For candidate percentages, the RMSE is 3.1-3.6 points on average, excluding the idiosyncratic contests (Alaska, Hawaii and D.C.), compared with 2.2-3.5 points in 2016. As a point of comparison, the RMSE of Obama's vote share from top forecasting models in 2012 – when such models performed exceptionally well – ranged from 1.9 to 2.8 (Muelhauser and Branwen 2012). Coverage rates of the 95 percent intervals of the candidate vote share estimates reach 90.2 % on average across the elections (excluding 2012), suggesting decent but not perfect model calibration.

Sharpening the 2000-2008 models to take greater account of contest-specific dynamics (known pre-election) may make them more accurate still, underlining the importance of incorporating prior knowledge into the models to best predict the eventual results. The greater accuracy of MRP in more recent contests may also reflect more early voters (McDonald 2016), greater predictive power of the included demographic variables for these elections and improved polling techniques, such as the inclusion of cell phones and Spanish-language interviewing.<sup>17</sup>

---

<sup>17</sup> Indeed, income, rather than education, is probably a better predictor of voting behavior in the earlier elections (Gelman et al. 2007). Including variables to account for candidate home state effects may have also improved the estimates, but were not relevant for the last two elections.

|                                 | 2000      | 2004      | 2008      | 2012     | 2016     | Average<br>(excluding<br>2012) |
|---------------------------------|-----------|-----------|-----------|----------|----------|--------------------------------|
| Democrat %                      | 46.9%     | 47.6%     | 54.9%     | 51.1%    | 46.8%    |                                |
| Republican %                    | 47.4%     | 48.6%     | 42.5%     | 46.1%    | 44.3%    |                                |
| Margin                          | -0.5 pts. | -1.1 pts. | 12.4 pts. | 5.0 pts. | 2.6 pts. |                                |
| Democrat EVs                    | 279       | 253       | 388       | 303      | 297      |                                |
| Republican EVs                  | 258       | 284       | 150       | 235      | 241      |                                |
| Democrat win probability        | 67.1%     | 21.9%     | 100.0%    | 100.0%   | 36%      |                                |
| Correct state predictions       | 46        | 46        | 47        | 49       | 50       | 47.6                           |
| RMSE margin all                 | 7.6%      | 6.8%      | 8.1%      | 4.6%     | 5.6%     | 7.03%                          |
| RMSE margin no AK, HI, DC       | 6.6%      | 5.8%      | 7.3%      | 4.5%     | 4.3%     | 6.01%                          |
| RMSE margin battlegrounds       | 4.8%      | 4.0%      | 5.2%      | 3.2%     | 2.7%     | 4.15%                          |
| RMSE Dem. % (no AK, HI, DC)     | 3.4%      | 3.4%      | 3.4%      | 2.2%     | 2.2%     | 3.11%                          |
| RMSE Rep. % (no AK, HI, DC)     | 3.7%      | 3.2%      | 4.0%      | 2.4%     | 3.5%     | 3.59%                          |
| Dem. 95% interval coverage rate | 90.2%     | 96.1%     | 84.3%     | 88.2%    | 90.2%    | 90.2%                          |
| Rep. 95% interval coverage rate | 94.1%     | 90.2%     | 82.4%     | 82.4%    | 90.2%    | 89.2%                          |

**Table 5.** Summary of MRP estimates from prior elections. Prior MRP models differed slightly in terms of the state-level predictors included, and some included cross-level interactions. Models were estimated on comparable Ns; for 2008 and 2012, this included the full tracking data, while for 2000 and 2004 the models were estimated on a subset of tracking data, given their larger samples and longer durations.

## VII. Discussion

While higher-quality polling in swing states likely would have improved predictions in the 2016 election, using two-step MRP models to estimate both turnout and vote preferences with national pre-election polls provides an attractive alternative. As this paper demonstrates, even with relatively small state-level sample sizes, this MRP approach substantially outperforms leading polling aggregators in the 2016 election, and analyses of previous elections support this approach.

Relative to state polling averages, this performance likely is related to factors including the quality of the underlying data and attributes specific to our approach. By estimating voter

poststratification targets using the same pre-election data rather than data from the prior election, we can account for consequential shifts in turnout among groups. Second, using a single national survey ensured that our estimates are based on data collected with the same methods across states, while state-level surveys averaged by aggregators vary widely in methods and quality. To the extent that lower-quality or poorly devised polling methods produce inaccurate estimates, the presumed canceling-out benefits of aggregation can lead to biased and misleading results. Relatedly, the analysis reported here is based on one of the most methodologically sound probability-based RDD surveys of its type in the country (Silver 2016), potentially an advantage over non-probability data.

Our MRP approach also offers an alternative to traditional survey weighting and likely voter modeling that overcomes some of the challenges faced by standard weighting techniques (Gelman 2007) – either iterative proportional fitting, which does not guarantee precise subgroup sizes, or cell weighting, which can be compromised by limited sample sizes. MRP is analogous in many ways to cell weighting, without the troubles associated with zero- or small- $n$  cells. In the analysis presented here, the model estimates were poststratified on 10,200 cells, essentially a much finer-grained weighting scheme than either rake or typical cell weighting. Further, by using Bayesian estimation to model turnout and vote preferences, we can account for uncertainty in both sets of quantities; most prior research has not fully incorporated uncertainty in turnout estimates.

Sample size efficiency gains represent an additional advantage of this approach. Across elections, our models produced RMSEs on the major party candidate vote shares typically ranging from 3 to 3.5 percentage points, which translates to state level margins of error of about 6-7 points. To achieve 7-point error margins with 51 state polls would require approximately 340

respondents per state or 17,340 respondents total,<sup>18</sup> compared with 9,485 general population respondents and 6,825 likely voters in the 2016 ABC/Post tracking poll.

However, MRP is not a cure-all for challenges facing pollsters. One important restriction is that it does not provide a single weight that can be used for all variables in a survey; it requires each outcome of interest to be modeled separately. Also, the accuracy of the approach is strongly related to the degree to which demographic variables available for poststratification predict voting behavior (Warshaw and Rodden 2012). While such demographic variables have been highly predictive in the past several elections, the future is unknown. To ensure continued accuracy, researchers employing the technique need to adjust demographic and state-level predictors included in the model to the dynamics of any given election, based on exploratory data analysis and other available information (Warshaw and Rodden 2012; Buttice and Highton 2013).

Future research may improve the accuracy of the MRP approach employed here. Other strategies for estimating turnout (e.g., models that combine estimates from CPS and pre-election data) could enhance subgroup-level turnout estimates, while oversampling lower population swing states could increase the likelihood of correctly predicting Electoral College outcomes. Future research also could examine whether and how to poststratify on variables such as partisan identification or past vote (Wang et al. 2014; YouGov 2016).

---

<sup>18</sup> While simple random sampling would imply sample sizes of approximately 200 per state to reach this level of accuracy, Chirani-Mehr, Rothschild, Goel and Gelman's (2016) analysis of state level presidential polling errors suggest design effects of approximately 1.73, substantially increasing required sample sizes.

For all its utility, MRP can only take us so far in understanding the dynamics of an election. While the method can produce precise state- and group-level turnout and candidate-support predictions, the key question of how groups come to their choices is best explored with substantive survey questions and analysis. That said, we recognize the intense media and public interest in discerning the likely winner of elections before they are held. If such predictions are to remain a dominant element of our pre-election landscape, it is best that they be accurate, an aim that we hope can be advanced by the approach described in this paper.

## References

- Bates, Douglas. 2005. "Fitting Linear Models in R Using the lme4 Package." *R News* 5(1):27–30.
- Buttice, Matthew K., and Benjamin Highton. 2013. "How does multilevel regression and poststratification perform with conventional national surveys?" *Political Analysis* 21(4): 449–467.
- Cohn, Nate. 2016. "There are More White Voters than People Think. That's Good News for Trump." *The New York Times*, June 9.  
<https://www.nytimes.com/2016/06/10/upshot/there-are-more-white-voters-than-people-think-thats-good-news-for-trump.html>
- Cohn, Nate and Amanda Cox. 2016. "The Voting Habits of Americans Like You." *The New York Times*, June 10. <https://www.nytimes.com/interactive/2016/06/10/upshot/voting-habits-turnout-partisanship.html>
- Cohn, Nate, Josh Katz and Kevin Quealy. 2016. "Putting the Polling Miss of the 2016 Election in Perspective." *The New York Times*, Nov. 13.  
<https://www.nytimes.com/interactive/2016/11/13/upshot/putting-the-polling-miss-of-2016-in-perspective.html>
- Eten, Harry. 2016. "Trump is Just a Normal Polling Error Behind Clinton." *FiveThirtyEight*, Nov. 4. <https://fivethirtyeight.com/features/trump-is-just-a-normal-polling-error-behind-clinton/>
- Gelman, Andrew. 2007. "Struggles with survey weighting and regression modeling." *Statistical Science* 22(2): 153–164.
- Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel-Hierarchical Models*. Cambridge: Cambridge University Press.
- Gelman, Andrew, and Thomas C. Little. 1997. "Poststratification into many categories using hierarchical logistic regression." *Survey Methodology* 23 (2): 127–35.
- Ghitza, Yair, and Andrew Gelman. 2013. "Deep interactions with MRP: Election turnout and voting patterns among small electoral subgroups." *American Journal of Political Science* 57(3): 762–776.
- Hur, Aram and Christopher Achen. 2013. "Coding Voter Turnout Responses in the Current Population Survey." *Public Opinion Quarterly* 77(4): 985–993.
- Lauderdale, Benjamin. 2016. "Election Model Methodology." *YouGov*, Oct. 3.  
<https://today.yougov.com/news/2016/10/03/election-model-methodology/>
- Lax, Jeffrey R. and Justin H. Phillips. 2009a. "Gay Rights in the States: Public Opinion and Policy Responsiveness." *American Political Science Review* 103(3):367–86.
- Lax, Jeffrey R. and Justin H. Phillips. 2009b. "How Should We Estimate Public Opinion in the States?" *American Journal of Political Science* 53(1):107–21.



- McDonald, Michael P. 2007. "The True Electorate: A Cross-Validation of Voter Registration Files and Election Survey Demographics." *Public Opinion Quarterly* 71 (4): 588-602.
- McDonald, Michael P. 2016. Voter Turnout Data. *United States Election Project*.  
<http://www.electproject.org/home/voter-turnout/voter-turnout-data>
- McDonald, Michael P. 2017. "CPS Vote Over-Report and Non-Response Bias Correction."  
Available at <http://www.electproject.org/home/voter-turnout/cps-methodology>.
- Mercer, Andrew, Claudia Deane and Kyley McGeeney. 2016. "Why 2016 election polls missed their mark." Pew Research Center, Nov. 9. <http://www.pewresearch.org/fact-tank/2016/11/09/why-2016-election-polls-missed-their-mark/>
- Park, David K., Andrew Gelman and Joseph Bafumi. 2004. "Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls." *Political Analysis* 12(4):375–85.
- Pacheco, Julianna. 2011. "Using National Surveys to Measure State Public Opinion over Time: A Guideline for Scholars and an Application." *State Politics and Policy Quarterly* 11(4): 415–39.
- Sanchez, Gabriel and Matt A. Barreto. 2016. "In record numbers, Latinos voted overwhelmingly against Trump. We did the research." *Washington Post*, Nov. 11.  
<https://www.washingtonpost.com/news/monkey-cage/wp/2016/11/11/in-record-numbers-latinos-voted-overwhelmingly-against-trump-we-did-the-research/>
- Shirani-Mehr, Houshmand, David Rothschild, Sharad Goel and Andrew Gelman. 2016. "Disentangling Bias and Variance in Election Polls."  
<http://researchdmr.com/ProbabilityTotalError.pdf>
- Silver, Nate. "Pollsters Probably Didn't Talk to Enough White Voters without College Degrees." 2016. *FiveThirtyEight*, Dec. 1. <https://fivethirtyeight.com/features/pollsters-probably-didnt-talk-to-enough-white-voters-without-college-degrees/>
- Silver, Nate. 2017. "The Real Story of 2016." *FiveThirtyEight*, Jan. 19.  
<http://fivethirtyeight.com/features/the-real-story-of-2016/>
- Stan Development Team. 2016. *Stan Modeling Language Users Guide and Reference Manual*.  
<http://mc-stan.org/documentation/>
- Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman. 2014. "Forecasting elections with non-representative polls." *International Journal of Forecasting* 31(3): 980-991.
- Warshaw, Christopher and Jonathan Rodden. 2012. "How Should We Measure District-Level Public Opinion on Individual Issues?" *Journal of Politics* 74(1): 203–219.

## **Appendix A: Turnout**

The first step of the analysis is to estimate turnout by demographic group. Table A1 shows the results of two turnout models using the 18-wave tracking dataset as well as the 2012 CPS Voter Supplement Survey. The first ABC/Post-based model classifies voters as 1 if they said they had voted or definitely would vote, otherwise 0. In the second, the dependent variable is restricted only to include voters in the first model who also reported voting in the 2012 presidential election. For the 2012 CPS-based models, those who reported voting are coded as 1 and those saying they did not vote 0, following the code provided by McDonald (2017). The second, more restrictive operationalization of turnout intent in the ABC/Post data proved most predictive of actual turnout in 2012 (as it did in 2016); we use this model to predict turnout poststratification targets in for our main models.

The subgroup-level variables that proved most predictive of turnout are education, age, and (at some distance) gender and race. Variance in turnout also is explained, but to lesser extents, by state of residence and by interactions between race and the other demographic variables. As far as fixed effects, the turnout in the respondent's state in 2012 is highly predictive in the main ABC/Post model; while positive, the coefficient associated with living in a battleground state does not reach conventional statistical significance. Overall, these findings are consistent with research linking age and socioeconomic status with turnout.

**Table A1: Predicting voter turnout, multilevel logistic regressions**

|   | 2016<br>ABC/Post:<br>Definite<br>voter | 2016<br>ABC/Post:<br>Definite<br>voter &<br>voted in 2012 | 2012 CPS<br>Reported<br>voter |
|---|--|---|-------------------------------|
| <b><i>Fixed part (<math>\beta</math>)</i></b>   |  |   |                               |
| Intercept                                       | 0.05<br>(0.86)                         | -0.62<br>(0.96)   | -2.29**<br>(0.53)             |
| VAP/VEP vote 2012                               | 0.91<br>(0.60)                         | 1.48**<br>(0.62)  | 4.51**<br>(0.31)              |
| Battleground                                    | 0.11<br>(0.08)                         | 0.08<br>(0.09)  |                               |
| <b><i>Random part (<math>\sigma</math>)</i></b> |  |   |                               |
| Gender  | 0.71                                   | 0.74  | 0.13                          |
| Age   | 0.93                                   | 1.16  | 0.50                          |
| Race  | 0.60                                   | 0.61  | 0.41                          |
| Education                                       | 1.20                                   | 1.33  | 0.85                          |
| State   | 0.10                                   | 0.14  | 0.06                          |
| Region  |  |   |                               |
| AgeXedu   | 0.16                                   | 0.15  | 0.11                          |
| GenderXedu                                      |  |   |                               |
| RaceXage  | 0.21                                   | 0.16  | 0.11                          |
| RaceXedu  | 0.15                                   | 0.10  | 0.22                          |
| RaceXgender                                     | 0.24                                   | 0.22  | 0.16                          |
| RaceXregion                                     | 0.14                                   | 0.13  | 0.20                          |
| Survey wave                                     | 0.17                                   | 0.21  |                               |
| N   | 9024                                   | 9024  | 82820                         |

\*\*p<0.01 p<0.05 (fixed part)

Estimates using the 2016 ABC/Post tracking survey are from Bayesian multilevel logistic regressions estimated using Stan through the rstanarm package in R, which implements MCMC sampling of the posterior distributions. Four chains of 2,000 samples were run, including 1,000 warm-up samples. Parameter estimates are posterior medians, with standard deviations in parentheses. The dependent variable for the first column is equal to 1 if the respondent said s/he was registered to vote and definitely would vote or had already voted, otherwise 0. The dependent variable in the second model is the same as the first except likely voters (1) also said they voted in the 2012 presidential election. The CPS-based models are multilevel logistic regressions estimated via maximum likelihood. The fixed part reports unstandardized beta coefficients along with standard errors, while the random part reports the estimated standard deviations of the random effects. Random effects with zero estimated variances were removed from the models.

**Table A2: State-by-State Turnout Predictions, Alternative MRP Models**

|                | 2016 ABC/Post MRP Predictions |                      |                             |      | 2012 CPS MRP Predictions |            |
|----------------|-------------------------------|----------------------|-----------------------------|------|--------------------------|------------|
|                | VAP<br>Turnout                | Median<br>Prediction | 95%<br>Credible<br>Interval |      | VEP<br>Turnout           | Prediction |
| National       | 54.7                          | 60.9                 | 59.0                        | 62.7 | 59.3                     | 60.8       |
| Alabama        | 56.3                          | 62.3                 | 57.7                        | 67.4 | 59.0                     | 60.7       |
| Alaska         | 57.4                          | 58.0                 | 52.2                        | 63.8 | 61.3                     | 56.9       |
| Arizona        | 49.2                          | 57.8                 | 53.5                        | 62.2 | 55.0                     | 54.6       |
| Arkansas       | 49.4                          | 57.5                 | 52.0                        | 62.1 | 52.8                     | 52.8       |
| California     | 47.0                          | 53.8                 | 50.9                        | 56.8 | 56.7                     | 56.0       |
| Colorado       | 64.6                          | 66.4                 | 61.8                        | 71.0 | 70.1                     | 71.7       |
| Connecticut    | 58.3                          | 66.0                 | 61.5                        | 70.8 | 64.2                     | 63.2       |
| Delaware       | 59.2                          | 63.5                 | 57.6                        | 68.6 | 64.4                     | 63.6       |
| DC             | 55.4                          | 63.0                 | 57.5                        | 69.3 | 60.9                     | 69.5       |
| Florida        | 56.9                          | 61.8                 | 58.5                        | 65.1 | 64.6                     | 64.3       |
| Georgia        | 52.6                          | 63.6                 | 59.4                        | 68.0 | 59.2                     | 61.5       |
| Hawaii         | 38.3                          | 51.5                 | 43.4                        | 58.5 | 42.2                     | 43.2       |
| Idaho          | 55.0                          | 61.7                 | 56.0                        | 66.8 | 59.1                     | 60.9       |
| Illinois       | 56.1                          | 61.6                 | 57.8                        | 65.4 | 61.9                     | 61.0       |
| Indiana        | 54.0                          | 60.5                 | 56.0                        | 64.5 | 56.4                     | 56.0       |
| Iowa           | 65.1                          | 69.1                 | 64.2                        | 74.4 | 68.4                     | 70.9       |
| Kansas         | 54.0                          | 62.1                 | 56.7                        | 66.8 | 57.7                     | 59.3       |
| Kentucky       | 56.1                          | 59.1                 | 53.7                        | 63.5 | 58.7                     | 56.8       |
| Louisiana      | 56.8                          | 62.6                 | 57.7                        | 68.7 | 60.0                     | 62.3       |
| Maine          | 69.4                          | 69.1                 | 63.2                        | 74.6 | 70.5                     | 69.9       |
| Maryland       | 59.5                          | 64.2                 | 59.6                        | 68.4 | 66.6                     | 68.8       |
| Massachusetts  | 61.1                          | 67.0                 | 62.8                        | 71.5 | 67.2                     | 66.9       |
| Michigan       | 62.0                          | 66.7                 | 62.5                        | 70.5 | 64.7                     | 66.4       |
| Minnesota      | 69.4                          | 70.2                 | 65.7                        | 74.5 | 74.2                     | 76.7       |
| Mississippi    | 53.3                          | 59.6                 | 53.7                        | 64.4 | 55.6                     | 62.3       |
| Missouri       | 59.6                          | 65.2                 | 60.3                        | 70.0 | 62.3                     | 63.7       |
| Montana        | 60.5                          | 65.8                 | 59.8                        | 71.5 | 61.8                     | 64.6       |
| Nebraska       | 58.8                          | 63.9                 | 58.2                        | 68.8 | 62.5                     | 62.7       |
| Nevada         | 49.4                          | 58.4                 | 53.2                        | 64.4 | 57.3                     | 56.7       |
| New Hampshire  | 69.1                          | 71.4                 | 66.0                        | 76.4 | 71.4                     | 72.3       |
| New Jersey     | 55.7                          | 61.4                 | 56.8                        | 65.3 | 64.1                     | 62.3       |
| New Mexico     | 50.2                          | 53.8                 | 48.6                        | 58.9 | 54.8                     | 54.1       |
| New York       | 49.6                          | 55.4                 | 51.5                        | 59.0 | 56.8                     | 55.1       |
| North Carolina | 60.2                          | 66.4                 | 62.5                        | 70.7 | 64.8                     | 65.5       |
| North Dakota   | 59.3                          | 63.9                 | 57.8                        | 69.0 | 60.9                     | 61.2       |
| Ohio           | 61.0                          | 66.7                 | 62.8                        | 70.4 | 62.9                     | 66.5       |
| Oklahoma       | 49.0                          | 57.3                 | 52.4                        | 61.9 | 52.4                     | 49.7       |
| Oregon         | 61.7                          | 66.3                 | 61.9                        | 71.0 | 66.4                     | 65.3       |
| Pennsylvania   | 60.5                          | 65.9                 | 62.0                        | 69.6 | 63.6                     | 60.4       |
| Rhode Island   | 54.7                          | 63.8                 | 58.5                        | 69.3 | 59.0                     | 58.4       |

|                |      |      |      |      |      |      |
|----------------|------|------|------|------|------|------|
| South Carolina | 54.2 | 64.5 | 59.8 | 69.5 | 56.7 | 59.4 |
| South Dakota   | 56.6 | 63.5 | 57.4 | 68.8 | 58.5 | 60.4 |
| Tennessee      | 48.6 | 60.1 | 55.6 | 64.4 | 51.2 | 54.3 |
| Texas          | 43.4 | 52.8 | 49.5 | 56.2 | 51.6 | 52.0 |
| Utah           | 52.8 | 58.3 | 52.7 | 63.2 | 56.7 | 57.2 |
| Vermont        | 62.3 | 66.5 | 59.9 | 71.4 | 63.7 | 64.3 |
| Virginia       | 60.8 | 62.5 | 57.1 | 66.8 | 66.1 | 67.1 |
| Washington     | 58.3 | 64.2 | 59.9 | 68.1 | 64.8 | 66.3 |
| West Virginia  | 49.1 | 57.0 | 49.9 | 62.2 | 50.1 | 46.6 |
| Wisconsin      | 66.2 | 68.3 | 63.5 | 72.4 | 69.4 | 73.5 |
| Wyoming        | 57.3 | 62.1 | 55.7 | 67.1 | 59.7 | 60.8 |

## **Appendix B: Preference Models**

The next set of models estimated preferences for Clinton (vs. other candidates) and Trump (vs. other candidates) using a similar set of demographic predictors at the individual level. Coefficient estimates are given in Table B1. State-by-state preference estimates based on models M3a and M3b with the two poststratification approaches are available in Table B2. As described in the main text, these models are highly accurate, missing the correct outcome only in one state. Predictions from the models without state-level covariates as well as the one with just prior vote shares do worse, highlighting the importance of including such variables. The models without any state-level variables correctly predicted only 45 states, while the ones only including 2012 candidate shares predicted 47 states.

**Table B1: Predicting voter support for Clinton and Trump, multilevel logistic regressions**

|   | Support Clinton |                  |                   | Support Trump   |                   |                   |
|---|-----------------|------------------|-------------------|-----------------|-------------------|-------------------|
|   | M1a             | M2a              | M3a               | M1b             | M2b               | M3b               |
| <b><i>Fixed part (<math>\beta</math>)</i></b>   |                 |                  |                   |                 |                   |                   |
| Intercept                                       | 0.54<br>(0.85)  | -1.38<br>(0.88)  | -1.05<br>(0.94)   | -1.29<br>(1.00) | -3.14**<br>(1.05) | -3.48**<br>(1.05) |
| Obama 2012 %                                    |                 | 3.81**<br>(0.55) | 3.82**<br>(0.60)  |                 |                   |                   |
| Romney 2012 %                                   |                 |                  |                   |                 | 3.89**<br>(0.63)  | 3.48**<br>(0.58)  |
| Black %   |                 |                  | -2.04**<br>(0.42) |                 |                   | 2.24**<br>(0.40)  |
| Hispanic %                                      |                 |                  | -0.12<br>(0.45)   |                 |                   | 0.75<br>(0.45)    |
| White evang. Prot. %                            |                 |                  | -0.07<br>(0.63)   |                 |                   | 0.88<br>(0.65)    |
| <b><i>Random part (<math>\sigma</math>)</i></b> |                 |                  |                   |                 |                   |                   |
| Gender  | 1.12            | 1.13             | 1.08              | 1.14            | 1.13              | 1.19              |
| Age   | 0.13            | 0.13             | 0.12              | 0.43            | 0.44              | 0.43              |
| Race  | 1.38            | 1.33             | 1.37              | 1.55            | 1.53              | 1.58              |
| Education                                       | 0.33            | 0.32             | 0.33              | 0.37            | 0.36              | 0.38              |
| State   | 0.29            | 0.18             | 0.11              | 0.31            | 0.21              | 0.11              |
| Region  | 0.36            | 0.14             |                   | 0.40            | 0.16              |                   |
| AgeXedu   |                 |                  |                   | 0.11            | 0.12              | 0.11              |
| GenderXedu                                      |                 |                  |                   |                 |                   |                   |
| RaceXage  | 0.11            | 0.11             | 0.11              |                 |                   |                   |
| RaceXedu  | 0.36            | 0.35             | 0.36              | 0.41            | 0.39              | 0.40              |
| RaceXgender                                     | 0.42            | 0.43             | 0.40              | 0.89            | 0.85              | 0.87              |
| RaceXregion                                     |                 |                  |                   |                 |                   |                   |
| Survey wave                                     | 0.08            | 0.08             | 0.08              | 0.15            | 0.16              | 0.16              |
| N   | 6376            | 6376             | 6376              | 6376            | 6376              | 6376              |

\*\*p<0.01 \*p<0.05 †p<0.10 (fixed part)

Estimates using the 2016 ABC/Post tracking survey are from Bayesian multilevel logistic regressions estimated using Stan through the rstanarm package in R, which implements MCMC sampling of the posterior distributions. Four chains of 2,000 samples were run, including 1,000 warm-up samples. Parameter estimates are posterior medians, with standard deviations in parentheses.

**Table B2: ABC/Post MRP State-by-State Model Predictions, ABC/Post-based turnout and preference models**

|                | Clinton Vote Share |                   |                       |      | Trump Vote Share |                   |                       |      | Clinton win probability |
|----------------|--------------------|-------------------|-----------------------|------|------------------|-------------------|-----------------------|------|-------------------------|
|                | Actual             | Median Prediction | 95% Credible Interval |      | Actual           | Median Prediction | 95% Credible Interval |      |                         |
| National vote  | 48.2               | 46.8              | 45.2                  | 48.5 | 46.1             | 44.3              | 42.2                  | 46.2 |                         |
| Alabama        | 34.4               | 34.4              | 28.7                  | 38.4 | 62.1             | 59.5              | 55.1                  | 65.5 | 0%                      |
| Alaska         | 36.6               | 38.3              | 31.6                  | 44.1 | 51.3             | 47.4              | 41.4                  | 54.6 | 2%                      |
| Arizona        | 45.1               | 42.3              | 37.6                  | 46.6 | 48.7             | 48.7              | 44.1                  | 54.0 | 2%                      |
| Arkansas       | 33.7               | 33.1              | 29.2                  | 38.2 | 60.6             | 59.5              | 53.6                  | 64.0 | 0%                      |
| California     | 61.7               | 58.6              | 55.2                  | 61.9 | 31.6             | 32.7              | 29.5                  | 36.0 | 100%                    |
| Colorado       | 48.2               | 48.2              | 43.2                  | 52.8 | 43.3             | 41.2              | 35.9                  | 45.9 | 98%                     |
| Connecticut    | 54.6               | 53.4              | 49.1                  | 58.6 | 40.9             | 37.3              | 32.8                  | 42.4 | 100%                    |
| Delaware       | 53.4               | 50.9              | 45.0                  | 55.4 | 41.9             | 40.8              | 36.6                  | 47.1 | 99%                     |
| DC             | 90.9               | 80.1              | 73.1                  | 86.4 | 4.1              | 13.4              | 8.9                   | 19.1 | 100%                    |
| Florida        | 47.8               | 46.4              | 43.2                  | 49.8 | 49.0             | 46.8              | 43.5                  | 50.6 | 44%                     |
| Georgia        | 45.9               | 43.4              | 39.7                  | 47.6 | 51.0             | 47.9              | 43.5                  | 51.8 | 7%                      |
| Hawaii         | 62.2               | 70.1              | 63.9                  | 76.8 | 30.0             | 21.1              | 15.8                  | 26.5 | 100%                    |
| Idaho          | 27.5               | 28.6              | 22.8                  | 34.5 | 59.3             | 59.8              | 52.9                  | 67.5 | 0%                      |
| Illinois       | 55.8               | 52.7              | 49.3                  | 57.3 | 38.8             | 38.0              | 33.4                  | 41.7 | 100%                    |
| Indiana        | 37.8               | 38.1              | 34.6                  | 43.3 | 56.9             | 52.6              | 47.3                  | 56.6 | 0%                      |
| Iowa           | 41.7               | 42.4              | 36.3                  | 47.6 | 51.1             | 48.4              | 43.4                  | 55.2 | 5%                      |
| Kansas         | 36.1               | 34.8              | 30.5                  | 40.5 | 56.7             | 55.3              | 49.7                  | 60.6 | 0%                      |
| Kentucky       | 32.7               | 33.3              | 29.4                  | 38.8 | 62.5             | 59.2              | 53.7                  | 63.8 | 0%                      |
| Louisiana      | 38.4               | 38.0              | 33.0                  | 42.2 | 58.1             | 53.1              | 48.7                  | 58.2 | 0%                      |
| Maine          | 47.8               | 47.3              | 41.3                  | 53.1 | 44.9             | 41.3              | 35.1                  | 47.0 | 93%                     |
| Maryland       | 60.3               | 57.6              | 53.3                  | 62.5 | 33.9             | 32.6              | 27.8                  | 36.9 | 100%                    |
| Massachusetts  | 60.0               | 55.9              | 51.4                  | 60.9 | 32.8             | 33.1              | 28.6                  | 37.8 | 100%                    |
| Michigan       | 47.3               | 45.7              | 41.7                  | 49.4 | 47.5             | 45.7              | 41.9                  | 50.0 | 49%                     |
| Minnesota      | 46.4               | 46.5              | 42.2                  | 51.9 | 44.9             | 41.4              | 35.3                  | 45.7 | 95%                     |
| Mississippi    | 40.1               | 39.9              | 34.0                  | 44.3 | 57.9             | 52.9              | 48.3                  | 58.5 | 0%                      |
| Missouri       | 38.1               | 37.7              | 33.1                  | 41.5 | 56.8             | 53.1              | 48.7                  | 57.9 | 0%                      |
| Montana        | 35.7               | 35.6              | 29.9                  | 40.9 | 56.2             | 53.1              | 47.4                  | 60.5 | 0%                      |
| Nebraska       | 33.7               | 33.8              | 28.7                  | 39.4 | 58.7             | 54.9              | 48.3                  | 61.0 | 0%                      |
| Nevada         | 47.9               | 48.5              | 43.4                  | 53.2 | 45.5             | 41.8              | 36.7                  | 47.1 | 97%                     |
| New Hampshire  | 46.8               | 45.8              | 40.2                  | 53.0 | 46.5             | 42.4              | 36.0                  | 48.6 | 81%                     |
| New Jersey     | 55.5               | 55.4              | 51.5                  | 61.1 | 41.4             | 35.3              | 30.2                  | 39.2 | 100%                    |
| New Mexico     | 48.3               | 53.9              | 45.7                  | 61.0 | 40.0             | 38.0              | 31.1                  | 45.4 | 100%                    |
| New York       | 59.0               | 57.3              | 52.6                  | 60.7 | 36.5             | 33.7              | 30.1                  | 37.9 | 100%                    |
| North Carolina | 46.2               | 42.6              | 38.4                  | 46.1 | 49.8             | 49.2              | 45.3                  | 53.3 | 1%                      |
| North Dakota   | 27.2               | 33.0              | 27.8                  | 38.6 | 63.0             | 54.7              | 48.4                  | 60.6 | 0%                      |
| Ohio           | 43.6               | 42.7              | 39.1                  | 46.2 | 51.7             | 48.0              | 44.2                  | 52.2 | 2%                      |
| Oklahoma       | 28.9               | 30.8              | 26.5                  | 35.0 | 65.3             | 61.8              | 56.8                  | 67.2 | 0%                      |
| Oregon         | 50.1               | 48.0              | 43.1                  | 53.1 | 39.1             | 40.3              | 35.1                  | 45.5 | 98%                     |
| Pennsylvania   | 47.9               | 43.7              | 39.5                  | 47.0 | 48.6             | 47.5              | 43.9                  | 52.5 | 7%                      |



|                |      |      |      |      |      |      |      |      |      |
|----------------|------|------|------|------|------|------|------|------|------|
| Rhode Island   | 54.4 | 55.5 | 48.8 | 60.8 | 38.9 | 33.6 | 28.6 | 39.4 | 100% |
| South Carolina | 40.7 | 39.7 | 35.6 | 43.9 | 54.9 | 52.7 | 48.4 | 56.8 | 0%   |
| South Dakota   | 31.7 | 33.7 | 28.1 | 38.9 | 61.5 | 55.4 | 49.6 | 62.0 | 0%   |
| Tennessee      | 34.7 | 35.5 | 31.6 | 40.8 | 60.7 | 57.8 | 52.4 | 62.1 | 0%   |
| Texas          | 43.2 | 43.1 | 39.2 | 47.2 | 52.2 | 49.2 | 45.1 | 53.0 | 2%   |
| Utah           | 27.5 | 31.0 | 23.5 | 38.5 | 45.5 | 52.3 | 44.0 | 60.4 | 0%   |
| Vermont        | 56.7 | 59.2 | 53.1 | 66.6 | 30.3 | 29.9 | 23.8 | 35.5 | 100% |
| Virginia       | 49.8 | 46.7 | 42.7 | 50.4 | 44.4 | 43.7 | 39.7 | 47.6 | 86%  |
| Washington     | 51.8 | 51.2 | 47.2 | 56.4 | 36.3 | 39.1 | 34.7 | 43.6 | 100% |
| West Virginia  | 26.5 | 28.7 | 23.0 | 33.8 | 68.6 | 64.3 | 58.4 | 70.1 | 0%   |
| Wisconsin      | 46.5 | 45.0 | 40.8 | 50.1 | 47.2 | 44.4 | 38.7 | 48.7 | 59%  |
| Wyoming        | 21.9 | 25.2 | 18.8 | 31.7 | 68.2 | 62.6 | 55.1 | 70.1 | 0%   |

## **Appendix C: Response Rates, Methodological Details and Question Wording**

The 2016 ABC News and ABC News/Washington Post Tracking poll was conducted from October 20, 2016 through November 7, 2016 using random-digit dialing to landlines and cell phones in the United States. A total of 9,930 respondents were interviewed in English or Spanish. The response rate (AAPOR RR3) for the full sample (landline and cell phone) was 14.7 percent (14.5 percent for cell phone, 17.7 percent for landline.)

The 2012 ABC News/Washington Post was conducted from October 18, 2012 through November 5, 2012 using random-digit dialing to landlines and cell phones in the United States. A total of 9,837 respondents were interviewed in English or Spanish. The response rate (AAPOR RR3) for the full sample (landline and cell phone) was 24.6 percent (27.7 percent for cell phone, 23.6 percent for landline.

The 2008 ABC News/Washington Post tracking poll was conducted from October 31, 2008 through November 3, 2008 using random-digit dialing to landlines and cell phones in the United States. A total of 10,213 respondents were interviewed in English or Spanish. The response rate (AAPOR RR3) for the full sample (landline and cell phone) was 34.1 percent (20.5 percent for cell phone, 36.5 percent for landline.)

The 2004 ABC News/Washington Post tracking poll was conducted from October 1, 2004 through November 1, 2004 using random-digit dialing to landlines in the United States. A total of 21,265 respondents were interviewed in English. The response rate (AAPOR RR3) for the full landline sample was 37.1 percent.

The 2000 ABC News/Washington Post tracking poll was conducted from October 12, 2000 through November 6, 2000 using random-digit dialing to landlines in the United States. A total of 16,721 respondents were interviewed in English. Response rates are not available.

Question wordings follow. Wordings used are same in all years unless noted.

**Registered voter:** Are you registered to vote at your present address, or not?

**Vote certainty:** I'd like you to rate the chances that you will vote in the presidential election\*:

Are you absolutely certain to vote, will you probably vote, are the chances 50-50 or less than that?

*\*"In November" used during October fielding in 2016, 2012, 2004, 2000. "Tomorrow" used on last day of fielding in 2008. Already voted accepted as volunteered answer in 2016, 2012, 2008, 2004.*

**Table C1: Candidate Preference Question Wording**

|      | <b>If already voted...</b>  | <b>All others</b>   |
|------|---|---|
| 2016 | <p>Confidentially and for statistical purposes only, did you vote for Hillary Clinton and Tim Kaine, the Democrats, Donald Trump and Mike Pence, the Republicans, Gary Johnson and Bill Weld of the Libertarian Party or Jill Stein and Ajamu Baraka of the Green Party?</p> <p>[IF REFUSED] We understand and respect your privacy. We're only asking for research purposes. All your answers are confidential. You can just tell me the number: Did you vote for ONE Hillary Clinton and Tim Kaine, TWO for Donald Trump and Mike Pence, THREE for Gary Johnson and Bill Weld, or FOUR for Jill Stein and Ajamu Baraka?</p> | <p>If the presidential election were being held today and the candidates were Hillary Clinton and Tim Kaine, the Democrats, Donald Trump and Mike Pence, the Republicans, Gary Johnson and Bill Weld of the Libertarian Party and Jill Stein and Ajamu Baraka of the Green Party, for whom would you vote?</p> <p>[IF NOT CLINTON, TRUMP, JOHNSON, STEIN OR WOULD NOT VOTE] Would you lean toward Clinton and Kaine, Trump and Pence, Johnson and Weld or Stein and Baraka?</p> |

|      |   |   |
|------|---|---|
| 2012 | Confidentially and for statistical purposes only, did you vote for Barack Obama and Joe Biden or Mitt Romney and Paul Ryan?       | <p>If the presidential election were being held today and the candidates were Barack Obama and Joe Biden, the Democrats and Mitt Romney and Paul Ryan, the Republicans, for whom would you vote?</p> <p>[IF NOT OBAMA, ROMNEY OR WOULD NOT VOTE] Which candidates are you leaning toward, Barack Obama and Joe Biden, the Democrats or Mitt Romney and Paul Ryan, the Republicans?</p>  |
| 2008 | Confidentially and for statistical purposes only, did you vote for Barack Obama and Joe Biden or for John McCain and Sarah Palin? | <p>If the 2008 presidential election were being held today and the candidates were Barack Obama and Joe Biden, the Democrats and John McCain and Sarah Palin, the Republicans, for whom would you vote?</p> <p>[IF NOT OBAMA, MCCAIN OR WOULD NOT VOTE] Which candidates are you leaning toward, Barack Obama and Joe Biden, the Democrats or John McCain and Sarah Palin, the Republicans?</p>   |
| 2004 | Which candidate did you vote for in the 2004 presidential election?   | <p>If the 2004 presidential election were being held today, would you vote for:<br/>[IN STATES WITH NADER ON BALLOT] George W. Bush and Dick Cheney, the Republicans, John Kerry and John Edwards, the Democrats, or Ralph Nader and Peter Camejo, the independents?</p> <p>[IN STATES WITHOUT NADER ON BALLOT] George W. Bush and Dick Cheney, the Republicans or John Kerry and John Edwards, the Democrats?</p> <p>[IF NOT BUSH, KERRY, NADER OR WOULD NOT VOTE] Which candidate are you leaning toward?</p> |

|      |                   |   |
|------|-------------------|---|
| 2000 | <i>Not asked.</i> | <p>The candidates in November's presidential election are Al Gore and Joseph Lieberman, the Democrats, George W. Bush and Dick Cheney, the Republicans, Ralph Nader and Winona LaDuke of the Green Party and Pat Buchanan and Ezola Foster of the Reform Party. If the election were being held today, who would you vote for - Gore, Bush, Nader or Buchanan?</p> <p>[IF NOT GORE, BUSH, NADER, BUCHANAN OR WOULD NOT VOTE] Who would you lean toward Gore, Bush, Nader or Buchanan?</p> |
|------|-------------------|---|

**Gender (2016 only):** Pardon but I'm required to verify - are you male or female? [IF NOT MALE/FEMALE] If you had to pick, would you say male or female?

**Age:** What is your age? [IF REFUSED]\* Could you please tell me if you are between the ages of 18 to 29, 30 to 39, 40 to 49, 50 to 64 or 65 or older?

*\*Refusal categories used in 2016, 2012, 2008.*

**Race and Ethnicity:** Are you of Hispanic origin or background? [IF YES] Are you White Hispanic or Black Hispanic? [IF NO] Are you white, black, or some other race?

**Education:** What was the last grade of school you completed?

