USCDornsife
Center for Economic and Social Research

The RAND Continuous 2012 Presidential Election Poll

Tania L. Gutsche, Arie Kapteyn, Erik Meijer, Bas Weerman

CESR WORKING PAPER SERIES

Paper Number: 2013-003   **cesr.usc.edu**

The RAND Continuous 2012 Presidential Election Poll

Tania L. Gutsche, Arie Kapteyn, Erik Meijer, Bas Weerman

University of Southern California and RAND Corporation

March 15, 2013

[Text from abstract up to and including the discussion is 6090 words]

Abstract

The RAND Continuous 2012 Presidential Election Poll (CPEP) was conducted within the American Life Panel, an internet panel recruited through traditional probability sampling to ensure representativeness. Because it asks the same respondents repeatedly about their voting preferences, observed changes are due to individuals changing their minds and not to random sampling fluctuations. The CPEP asks respondents to state both their preferences for a candidate and the likelihood that they will vote as a percent chance. Moreover, we asked the respondents' actual voting after the election, so we can study the predictive power both within sample and out of sample (the national results). The CPEP appears to have predicted well. Our final prediction of the difference in popular vote between Obama and Romney differed about half a percentage point from the final tally, which would place it near or at the top of the polling firms. The probabilistic questions, even months before the election, were strongly related to actual voting behavior. Our approach allows us to gain insights in stability of voting preferences and the effect of events on individual preferences; for example, various shifts can be clearly related to major events. The American Life Panel has a wealth of background characteristics that can be related to voting preferences.

1. Introduction

The RAND Continuous 2012 Presidential Election Poll implemented an innovative way to forecast the results of the U.S. Presidential election that was held on November 6, 2012. Its two main innovations are that respondents were asked to express the percent chance of voting for each candidate and that the study is conducted within a panel. The innovations have been pioneered by Delavande and Manski (2010) and they have shown this to be a successful method of eliciting respondents' likely voting behavior. The current study expands on this by implementing the method in a full-scale continuous election poll, with automatic daily postings of the latest forecasts. In order to arrive at these automatic daily forecasts, a complex weighting method needed to be implemented.

The contributions of this paper are the following. First, we partially replicate the validation of the probabilistic questions designed by Delavande and Manski (2010), for this new election season. Second, we show that the approach can be scaled up to a successful continuous election poll. Third, we relate changes in the forecasts to events in the campaign season and show that by and large the changes in forecasts match expectations. Fourth, we present results on average changes at the individual level in addition to changes in the aggregate forecasts, although as we indicate, understanding these individual-level changes requires further research. Fifth, we compare the forecasts based on the individuals' reports on their own voting behavior with the average chances that each candidate will win the presidency as stated by the respondents. Sixth, we relate political preferences and aggregate changes therein to individual characteristics of respondents.

The outline of the paper is as follows. Section 2 describes the methodology of the CPEP: the data source and sample sizes, the probabilistic questions and their predictive validity, the weighting method, and the implementation. The results, in the form of graphs as posted online during the election season, are shown and analyzed in Section 3. In the course of this, potential avenues of further research based on these data will be indicated. Section 4 summarizes and concludes.

2. Methodology

2.1 Participants and sample sizes

The sample for the study consists of members of the RAND American Life Panel (ALP). The ALP is an internet panel that is representative of the U.S. population of 18 years and older. Representativeness is attained by using standard survey sampling methods such as random digit dialing and address based sampling. Potential respondents need not have internet access. RAND provides members who did not previously have internet with internet access. See https://mmicdata.rand.org/alp/ for more information about the sampling and composition of the ALP.

In June of 2012, all 5301 ALP members who are U.S. citizens were invited to participate in this study. Eventually, 4293 members agreed to participate. Figure 1 shows the cumulative number of participants by date. On July 28, 3667 members had started participating. This is the sample we used for the daily forecasts and for this paper (the "analytic sample"). The rationale behind freezing the sample was that we wanted to make sure that changes we saw were actual changes in voting intentions and not due to changes in composition of the sample. However, we did compare the results with the results from the (then) full sample on two occasions (October 1 and after the elections) and we found only small differences with the analytic sample. The largest differences were observed for the first few weeks of the survey (in July), where the sample was the same as the analytic sample, but the base weights (to be discussed below) were different. We conclude from this that the later participants differ slightly from the earlier ones on demographic characteristics, but there does not appear to be selection bias conditional on the demographics that are used in constructing the weights. The data for all participants are available for researchers at https://mmicdata.rand.org/alp/.

--- FIGURE 1 ABOUT HERE ---

Upon participation, panel members were invited once a week to answer a limited number of questions about the election (see below). They received $2 for each completed survey. Every day, one seventh of the panel members were asked to respond to the questions discussed below. Typically, respondents answered the survey on the day they were invited, but we allowed them to delay by up to six days. After seven days, they were invited for the next survey and they could not return to the previous one. Thus, respondents received their invitations always on the same day of the week. The first invitations were sent on July 5, 2012.

We constructed daily samples based on a "rolling window". Specifically, each daily sample is based on a seven-day window, consisting of the past seven days, counting from the day it is assembled. This balances the desire to follow the developments daily and to ensure high statistical precision by having as large a sample as possible, with a stable sample composition. In the figures, each sample is assigned to its last survey day.

It is possible that by irregular timing of interviews, a respondent answered two surveys within the same sample window. If so, both are included and for simplicity they are treated as stochastically independent observations. This may lead to a slight underestimation of standard errors, but because most respondents completed the survey on the day they were invited, this is a minor issue.

Figure 2 shows the sample sizes of the interviews completed on each day (within the analytic sample of 3667), and the resulting sample sizes for the seven-day windows. This shows that the number of interviews is stable over time and does not suffer from appreciable attrition. It is particularly noteworthy that there is no drop-off in late October and early November, when Gallup suspended polling because of outages caused by Hurricane Sandy

(http://www.abc2news.com/dpp/news/political/gallup-polling-shut-down-after-superstorm-sandy-hits-swing-

), especially in the New York-New Jersey area. During this period, response was stable in New York and only slightly lower in New Jersey, which does not measurably affect our results. There are two notable drops in sample size, however, each followed by a peak. This is the result of respondents delaying completing the survey in these two periods. The first delay was in the Labor Day weekend (Monday, September 3, was Labor Day), when we suspect a number of panel members were away from home, and the second was September 22-24, also a weekend.

--- FIGURE 2 ABOUT HERE ---

Because the ALP is composed of individuals recruited from multiple sources over the course of about 10 years, it is difficult to compute a precise standardized response rate. To obtain an indication of the overall response rate of this study, we follow AAPOR (2011) and Callegaro and DiSogra (2008). In early 2012, the average recruitment rate (RECR) in the ALP was 22.1%, and the average retention rate (RETR) was 72.1%. Attrition between 2011 and 2012 was about 7%, but the panel has also been considerably expanded in this period with about 2000 additional members. Therefore, using the estimates from early 2012 will be conservative. The participation rate for the analytic sample of our study is 3667/5301 = 69.2%. From Figure 2, it follows that a typical sample contained about 3000 interviews, varying from roughly 2000 to 3500, resulting in an average completion rate (COMR) of about 3000/5301 = 56.6%, varying from about 37.7% to 66.0%. Thus, a conservative estimate of the average cumulative response rate (CUMRR2) is 22.1% × 72.1% × 56.6% = 9.0%, with a range of 6.0%-10.5%. This is a bit higher than the example given in Callegaro and DiSogra (2008).

2.2 Subjective probabilities and their predictive power

In contrast with most election polls, which ask respondents what they would do if the elections were held "today", we asked the respondents to predict the chances of voting in the election, and the chances of voting for the different

candidates conditional on voting in the election, as a percentage from 0%-100%. This methodology was pioneered by Delavande and Manski (2010) for the 2008 presidential election.

For our purposes, the three main questions, asked weekly of every respondent, are the following:

Q1:   What is the percent chance that you will vote in the Presidential election?

Q2:   If you do vote in the election, what is the percent chance that you will vote for Obama? And for Romney? And for someone else? Please provide percent chances in the table below:

Q3:   What do you think is the percent chance that each of the candidates for president will win the election?

Q2 and Q3 were followed by a table with for each of the three alternatives (Obama, Romney, other), a box in which respondents were asked to provide a percentage, and a Total box, in which the software computed the sum of the three, in order to facilitate consistent answers (which add up to 100%). Especially in the first few weeks, however, we did not always succeed in enforcing consistency, and we improved the questionnaire as a result of this. The inconsistent surveys, which fortunately make up only a very small fraction of the total, were excluded from the analyses[1].

The order of Obama and Romney was randomized, but the order was the same for Q2 and Q3 within a single interview. All these questions were presented on a single screen. See Kapteyn et al. (2012) or

https://mmicdata.rand.org/alp/elections.php for the screen layout. In addition to these weekly questions, five additional questions were asked twice since September 24 which we are not using for this paper but were shown on the same screen. These questions asked for (1) the percent chance that the margin between Obama and Romney would be less than 2 percentage points; (2) a question about a "feeling thermometer" (how warm does one feel towards a candidate); (3) anticipated time and effort to go to the polling place and vote; (4) whether politics and government are too difficult to understand, if citizens have a duty to vote, if public officials care what people think, etc.; (5) if the country would be

---

[1] For instance, one respondent entered a percent chance of voting for Romney of 1780, a percent chance of voting for Obama of -2000, and a percent chance of voting for another candidate of 320. Thus the numbers add up to 100, but are clearly outside the permissible range. Afterwards we restricted percentages to positive numbers (and kept imposing the restriction that they add up to 100).

heading in the right direction if each of the candidates were elected. Furthermore, from October 15 onward, respondents were asked whether they had already voted (early voting) and then instructed to fill in 100% in the box of the candidate they had voted for.

Table 1 shows the distribution of the percentages for four selected survey weeks. Although a majority of the answers is either 0% or 100%, a non-negligible fraction (between a fifth and a third) is strictly between these extremes. As expected, this fraction decreases as the time until the election becomes shorter.

--- TABLE 1 ABOUT HERE ---

The respondents were also asked two preliminary questions about their voting behavior in the 2008 U.S. presidential elections: Question I1 asked whether the individual voted in the 2008 Presidential election, and if yes, I2 asked whether the individual voted for Obama, McCain, or another candidate. Once respondents had answered these questions, they were not asked again in subsequent interviews. This information is used in constructing weights, as described in section 2.3. For some respondents who did not answer these questions, we were able to use their reports about their voting in 2008 when asked directly after the elections in 2008. For respondents who answered these questions both in 2008 and in 2012, the correspondence is about 90% (about 95% for those who answered in 2012 that they voted for Obama or McCain in 2008 and a bit lower for those who in 2012 answered that in 2008 they had not voted or had voted for another candidate). From other ALP surveys (the quarterly "My Household"), we have information about demographic and socio-economic variables, which were used in construction of the weights and for computing results for subpopulations.

After the election, respondents were asked to answer a survey about their actual voting, which allows us to assess the predictive power of the subjective probabilities for the actual outcomes. This partly replicates the analyses Delavande

and Manski (2010) report for 2008. Table 2 shows the result of linear regressions in which the dependent variable is whether the individual voted (0 for no and 100 for yes, to facilitate interpretation of the coefficients; top panel) or whether the individual voted for Obama, conditional on voting (bottom two panels, analogous coding). These regressions only include respondents who had not voted (early) at the time they stated their subjective probabilities. If the subjective probabilities were perfect predictors, their coefficient should be 1 and all other coefficients 0. Given the large concentration at 0% and 100%, for which the predictive bias can only be positive and negative, respectively, we expect coefficients smaller than 1.

--- TABLE 2 ABOUT HERE ---

For the probability to vote, we indeed see that the coefficient is smaller than 1 in the earlier survey weeks. Looking at both the coefficient and the intercept shows that the bias in earlier weeks is largely due to individuals who reported a low likelihood of voting but voted anyway, rather than individuals who reported a high likelihood and did not vote. In late September and late October, the coefficient is very close to 1 and the intercept is close to 0 and not statistically significant. We have also estimated the model with the preference for Obama (the relevant Q2 variable) as an additional explanatory variable, but the coefficient of this variable was always close to 0 and not significant. Hence, we conclude that the predictive bias is not systematically related to candidate preference.

For the probability of voting for Obama (middle panel of Table 2), the coefficient is close to 1 in all weeks and the intercept is close to 0 throughout. In the final two periods (09/19-09/25 and 10/25-10/30) the coefficient of voting for Obama is not statistically different from one and the intercept is not significantly different from zero, thus indicating that there was no predictive bias in the later periods. When we include the subjective probability to vote as an additional regressor (bottom panel) in the equation explaining voting for Obama, it is statistically significant and negative. Individuals who reported a high subjective probability of voting were more likely to overestimate their probability of voting for Obama than individuals who reported a low probability of voting. However, the inclusion of this additional

variable did not affect the coefficient of the subjective probability of voting for Obama, which is still very close to 1, and it did not increase predictive power as measured by the R-squared. Thus, for the purposes of forecasting the election, it would not have added much, even if we had known the coefficients beforehand. Also note that the R-squared values are monotonically increasing when the period gets closer to the election date.

We have also re-estimated these models with added state dummies, following reports in the media about differential difficulties with voting in different states. Some of these coefficients were statistically significant, but again they did not materially affect the coefficient of interest, nor the R-squareds.

Summarizing, we conclude that the subjective probabilities are good predictors of actual voting, with only some bias in the probability to vote in earlier periods, which is not related to candidate preference and thus does not bias the election forecasts (especially the Obama-Romney comparison).

2.3 Weighting and standard errors

Weights were computed in two stages. The first stage, the *base weight*, is the usual weight that is computed for most surveys in the ALP. It is based on matching demographics and socio-economic information to the CPS. This was computed only once for our analytic sample of 3667 participants. The second stage of weighting was done for each seven-day sample separately. This consisted of reweighting the sample (post-stratification) such that sample voting by strata in 2008 matches known population voting in 2008. This is based on the premise that the best predictor of future voting behavior is past voting behavior, and that any discrepancies in composition with respect to past voting behavior thus are likely to give biased predictions of voting behavior in the 2012 election. As mentioned above, 2008 voting was asked in two preliminary questions, and filled in from a 2008 survey when necessary and possible. We classified each individual uniquely in one of the following groups (post-strata): (1) 18-21 years old (in 2012), not eligible to vote in 2008;

(2) 22+, did not vote in 2008; (3) voted for Obama in 2008; (4) voted for McCain in 2008; (5) voted for another candidate in 2008; (6) 22+, unknown voting behavior in 2008. The information about the relative sizes of strata 2-5 was obtained from the United States Election Project (McDonald, 2012) and the official election results from the Federal Election Commission (2009). The size of stratum 1 was estimated from the CPS and the size of stratum 6 was taken to be proportional to its size relative to strata 2-5 in the analytic sample after applying the base weights.

Because some respondents answered the questions about 2008 voting later, the stratum allocation was initially not constant. However, on August 18, as part of a streamlining operation in order to reduce computing time (see below), stratum allocation was also frozen, and our analyses use this stratum allocation and its resulting weights.

Standard errors for estimators based on poststratified weights were be obtained using the methods and formulas given in StataCorp (2009a, p. 1026; 2009b, pp. 50-51, 160-162). See Kapteyn et al. (2012) for numbers and formulas for the weighting and the standard errors.

2.4 Automation and web implementation

RAND set up a publicly accessible website for the election poll (https://mmicdata.rand.org/alp/index.php?page=election), with some introductory text, links to information about the ALP and the methodology document (Kapteyn et al., 2012), and five graphs of results (see section 3). The graphs were updated daily through the following process: participants in the study were sent an email message to invite them to complete the survey once a week on their assigned day of the week. They then logged in to the server and answered the questions online. Each night at 00:05am Pacific Daylight Time, a script was run on the server, which downloaded the data entered by respondents until midnight and ran a Stata program that computed various statistics, which were written into new data files in csv format. When the website was accessed, a script on the website read the csv files and

created graphs of these, which were embedded in the website. The csv files were also available for download from the website (and still are).

The whole process was fully automated and did not require daily maintenance. However, on several occasions, some changes were made to streamline the process, most notably on August 18. Before then, the results for all earlier days were recomputed along with the most recent day, and this became computationally intensive. On August 18, this was changed such that only the results for the seven-day sample ending on the previous calendar day were computed and added to the previously computed and stored results for the earlier days. As mentioned in section 2.3, part of this streamlining also meant that stratum assignments were not recomputed.

Judging by the number of email messages we received, the website was well visited. It was also regularly mentioned in the media, for example, CNN, USA Today, and especially the New York Times' FiveThirtyEight blog. The posted results were also included as input in FiveThirtyEight's own forecast model.

3. Results

In this section, we present the graphs that were also posted on the website as part of the continuous poll, but add some context. Specifically, we indicate the timing of the key events in the election: the announcement of Paul Ryan as the Republican candidate for Vice President (August 11; "Ryan"), the Republican National Convention at which Mitt Romney officially became the Republican candidate (August 30, "RNC"), the Democratic National Convention at which Barack Obama officially became the Democratic candidate (September 6, "DNC"), the presidential debates (October 3, 16, 22), and the vice-presidential debate (October 11). We also indicate July 13, when Barack Obama made the controversial "You didn't build that" remark ("Build") and September 17, when a video of Mitt Romney making his controversial "47

percent" remark ("47%") appeared on the Internet. These were identified as key events on the Wikipedia page http://en.wikipedia.org/wiki/2012_U.S._Presidential_Election.

3.1 Intention to vote

Figure 3 shows the respondents' stated intention to vote, separately for Obama and Romney supporters. This figure is an adapted version of one of the graphs that was posted daily on the website. Whether someone is an Obama or Romney "supporter" is defined by the answer to the question who they would vote for, conditional on voting. However, because this is a probabilistic question, we cannot always uniquely assign individuals to either camp. The graph solves this by applying Bayes' theorem to obtain the probability of voting conditional on candidate preference as a function of the reported probabilities of voting and of voting for each candidate conditional on voting at all. In this graph, and analogous ones discussed below, the shaded area is centered on the average of the two curves and has a width of 1.96 standard errors of the difference between the two. Thus, the difference between the two is statistically significant at the 5% level if the curves are outside the shaded area. This graph, and the following ones, use the sampling weights as discussed in section 2.3. Detailed formulas underlying this graph and the ones discussed below are given in Kapteyn et al. (2012).

--- FIGURE 3 ABOUT HERE ---

Most of the time, the two curves are roughly parallel, with an upward trend, indicating that both Obama and Romney supporters became more determined to vote when the election day came nearer. However, from the second presidential debate until the election, and less extremely between the first presidential debate and the vice-presidential debate, the curves tend to move in opposite directions. This can happen because of two reasons. The first is that individuals have stable preferences for the candidates but become more or less encouraged to vote. The second is that individuals who report a low probability of voting (or, less likely, high) change their candidate preference, thereby

changing the composition of the two samples underlying the curves. Of course, a combination of these effects is also possible. The first debate was generally considered to be won by Mitt Romney and the second by Barack Obama[2]. If winning a debate would spur more, rather than less, enthusiasm for voting among a candidate's own supporters, we would expect the opposite pattern from what we see in the graph. However, if winning a debate has the effect of primarily moving individuals who are less likely to vote from the camp of the "loser" to the camp of the "winner", we would expect to see the pattern in the graphs. Our microdata allow us to investigate these explanations in more detail, which we intend to do in follow-up research.

As shown in section 2, the subjective probabilities of voting are good predictors of actual probabilities of voting, especially in the later weeks. However, when we compute the (weighted) fraction of the respondents who actually voted, it comes out about a percentage point lower for Obama supporters than the last predicted value (the one for Romney is right on the mark). A potential explanation for this would be a compositional issue, in which individuals who are more likely to miss an interview during the pre-election period are also less likely to vote. In this case, the graphs in the pre-election period would underrepresent individuals who are less likely to vote. This also suggests that improvements in the forecasts would be possible by taking this into account in the weighting. We will investigate these issues in further research. It should also be noted that the fraction among our respondents who voted is substantially higher than the national turnout, which is estimated to be 58.7% (McDonald, 2013). This is not unique to our data. For example, the 2008 American National Election Survey has the same limitation.

3.2 Daily forecasts and election outcomes

The main graph posted daily was the graph that showed the forecasted fraction of the votes for Obama and Romney. An adapted version of this graph is shown in Figure 4. This graph takes the subjective probabilities of voting in the election

_____

[2] See http://en.wikipedia.org/wiki/United_States_presidential_election_debates,_2012 ; Also see
http://www.cnn.com/2012/11/06/politics/21-things-look-back/index.html

into account to estimate the fraction of the votes for the candidates and thus can be viewed as a refinement of the "likely voter" method that is used in some form or another by many polling organizations. An advantage of our probabilistic questions is that we do not have to use an (untested) model for who is likely to vote in the election.

--- FIGURE 4 ABOUT HERE ---

There are many interesting aspects of this graph. Most noteworthy is that, with a few short exceptions, Obama is systematically higher than Romney. This contrasts with most other election polls conducted, which showed a much tighter race and in some cases (in particular, Gallup) showed Romney leading. The media typically focus on the *spread*, that is, the difference between the two curves. Our final forecast of this was 3.32 percentage points, whereas the final results showed a spread of 3.85 percentage points (Federal Election Commission, 2013), and our forecast was one of the closest[3]. Note, however, that the levels of both Obama and Romney are a bit too low, and correspondingly (but not shown) our forecast of the fraction of the votes for other candidates was too high. It is also noteworthy that the difference between Obama and Romney is mostly within the (pointwise) confidence band, although close to the limit of it in the final days. This corresponds well with the widely acclaimed New York Times' FiveThirtyEight final prediction, which gave Obama a 90.9% of winning (Silver, 2012).

As seen in the figure, the short periods in which the forecast for Romney was about equal to the one for Obama follow shortly after Obama's controversial "You didn't build that" remark, a few days after the announcement of Paul Ryan as Mitt Romney's running mate, and immediately after the Republican National Convention. Also, Romney's score improved after the first presidential debate, which was widely considered[4] to be won by him, but also for a few days after the third presidential debate, which was considered won by Obama. Conversely, the difference between Obama

---

[3] [3] http://fivethirtyeight.blogs.nytimes.com/2012/11/10/which-polls-fared-best-and-worst-in-the-2012-presidential-race/ Note that at the date of the blog (November 10) many votes were not counted yet. In the final tally the gap between Obama and Romney widened by about 1.2 percentage points from the count mentioned.

[4] See http://en.wikipedia.org/wiki/United_States_presidential_election_debates,_2012

and Romney increased noticeably after the Democratic National Convention, after Romney's controversial "47 percent" remark appeared in the news. We do not see a substantial increase in the difference in the aftermath of Hurricane Sandy[5], which hit the Northeastern United States—in particular New Jersey and the New York City area—on October 29-30. Republican governor of New Jersey Chris Christie extensively praised President Obama for his handling of this disaster.

3.3 Changes

The panel nature of our data allows us to study changes in expected voting behavior at the individual level. For this purpose, not voting and voting for "another candidate" were combined, and changes in percentages were translated into "from-to" percentages. The website included a graph with the fractions of changes from Obama to Romney and vice versa, as depicted in Figure 5. For computing changes, the "previous" survey is the most recent survey answered before the one included in the sample. This will typically be the survey in the prior week, but because of missed surveys or irregular timing, this previous survey may have been earlier, or within the current sample window. Also, if the actual previous survey was filled out inconsistently, it is dropped from the sample (see section 2) and the "previous" survey for computational purposes is the most recent consistently answered one.

--- FIGURE 5 ABOUT HERE ---

The average movement in either direction was between 1 and 1.5 percent on a weekly basis. The differences between the two curves roughly corroborate the trends in Figure 4. However, they are often difficult to interpret on a day-by-day basis, because they operate with a larger lag, because the number for a given day reflects the seven-day window ending on that day, and in terms of changes this means it is a weighted averages of changes between 13 days earlier and 6 days

---

[5] See http://en.wikipedia.org/wiki/Hurricane_Sandy for stories and links

earlier up to changes between 7 days earlier and the last day. Hence, even in the absence of irregular timing that could make these distances even larger, the curves in the graph reflect movements in the past two weeks and thus, with the frequent changes in direction in Figure 4, they regularly seemingly contradict that figure.

A perhaps more intuitive way to exploit the panel nature of the data would be to select a few key events, like the ones indicated in the figures, and study to what extent the responses in the first interview after the event differ from the ones in the last interview before the event. We will study this in follow-up research.

3.4 Respondents' predictions of the winner

From July 11 onward, the interviews included the additional question about what the respondents believed the chances were of Obama, Romney, or another candidate to win the presidency (Q3; see section 2.2). The averages of these answers were also posted on the website. Figure 6 shows the results.

--- FIGURE 6 ABOUT HERE ---

Clearly, this graph is very different from the one reflecting the respondents' own expected behavior. Respondents did not (only) report their own preferences, but took their expectations about others' behavior into account. Throughout, the chances for Obama were considered much higher than for Romney, and the differences are highly statistically significant. The movements are related to the ones in Figure 4, but much more dampened, and Romney never came close to Obama.

3.5 Subpopulations

The participants in the Continuous Presidential Election Poll are panel members of the American Life Panel, and from the many other surveys these members responded to, we have a wealth of information about them. Hence, we can relate their political preferences, and changes therein, to these other characteristics. For the website, we computed breakdowns of Figures 4-6 by a number of demographic characteristics: sex, age, race-ethnicity, education, income, education, labor force participation, family income, and state (Florida, Ohio, Pennsylvania). Each day, one of these graphs was randomly selected for posting on the website.

These graphs largely show the expected patterns: women, younger individuals, minorities, high educated, unemployed, and low income individuals are relatively more likely to favor Obama and men, older individuals, whites, lower educated, working or retired, and middle income individuals are relatively more likely to favor Romney. Interestingly, though, the expectations about the chances of Obama and Romney to win the presidency show an advantage of Obama throughout, which corroborates that this question is interpreted differently than the question about one's own voting behavior.

Here, we present two of these breakdowns: the forecasted fraction of the votes by sex and age. Figure 7 confirms that men were more likely to favor Romney and women more likely to favor Obama. It also shows that the drop of Romney after the "47 percent" video was completely due to changes among men. Remarkably, after the second presidential debate, in which Mitt Romney made his remark about "binders full of women", which was perceived as negative towards women in the media, Romney suffered only a very brief drop among women and a steep surge after that.

--- FIGURE 7 ABOUT HERE ---

Figure 8 shows the breakdown by age. Most interesting here is the reaction to the "47 percent" video, in which the 47 percent of individuals who were deemed dependent on government included many Social Security beneficiaries 65 years and older. The reaction to this video does appear to be most prominent among this age group.

--- FIGURE 8 ABOUT HERE ---

4. Discussion

In this paper, we have given an overview of the RAND 2012 Continuous Presidential Election Poll (CPEP). This poll had several unique characteristics. First, it was conducted in an online panel, RAND's American Life Panel, which has the advantage that changes in the forecasts are due to changes in preferences and not due to random changes in composition of the sample. The flipside of this is that the random sampling error inherent in each survey is the same over time in our poll. The panel character also allows us to study changes at the individual level, and relate preferences and changes in preferences to a wealth of other information we have about the panel members.

A second unique characteristic of the CPEP is that it used probabilistic questions: respondents were asked to state their subjective probability of voting in the election, and their subjective probability of voting for each candidate, conditional on voting in the election. Although many respondents reported a 0% or 100% chance, some 20-30% of the sample reported a number strictly between 0 and 100. Probabilistic questions have several advantages: they obviate the need for a likely voter model and they allow us to pick up changes from, say, 60% to 70% while on the other hand not overestimating the effects of small changes from, say 45% to 55%.

Third, we used a complex weighting method, which took not only demographic information into account, but also reweighted each daily forecast sample by voting behavior in 2008. If, due to random fluctuations or selective nonresponse, we would have an overrepresentation of, say, individuals who voted for Obama in 2008, this would have likely resulted in a forecast that would overstate the fraction who would vote for Obama in 2012. Our weighting strategy corrects for this.

The poll was fully automated, with invitation emails for interviews generated daily, and results also automatically computed and posted on the website daily.

In terms of forecasting the difference in popular vote between Obama and Romney, the CPEP was one of the most accurate, and praised as such in the media. Many other polls showed a very close race or even a Romney lead.

We have shown how patterns in the forecast over time are related to events in the campaign season. Most of these correspond with expectations.

The microdata from the CPEP are available for free to registered researchers who would like to analyze them[6]. Because of the reasons mentioned, the dataset, combined with other surveys in the ALP, forms a treasure trove for understanding individuals' political preferences and how they translate into voting behavior. We have given several examples of follow-up analyses we intend to do, but we invite any researcher with an interest in understanding voter behavior to download the data and analyze them.

---

[6] https://mmicdata.rand.org/alp/

References

AAPOR (2011). *Standard definitions: Final depositions of case codes and outcome rates for surveys* (7th ed.). The American Association for Public Opinion Research. Available from http://www.aapor.org/Standard_Definitions3/3557.htm

Callegaro, Mario, & DiSogra, Charles (2008). Computing response metrics for online panels. *Public Opinion Quarterly*, *72*, 1008-1032.

Delavande, Adeline, & Manski, Charles F. (2010). Probabilistic polling and voting in the 2008 presidential election: Evidence from the American Life Panel. *Public Opinion Quarterly,* 74, 433-459.

Federal Election Commission (2009). *2008 Official presidential general election results*. Available from http://www.fec.gov/pubrec/fe2008/2008presgeresults.pdf (accessed August 13, 2012).

Federal Election Commission (2013). *Official 2012 presidential general election results*. Available from http://www.fec.gov/pubrec/fe2012/2012presgeresults.pdf (accessed March 8, 2013).

Kapteyn, Arie, Meijer, Erik, & Weerman, Bas (2012). *Methodology of the RAND Continuous 2012 Presidential Election Poll*. Working Paper No. WR-961. Santa Monica, CA: RAND Corporation.

McDonald, Michael P. (2012). 2008 General Election Turnout Rates. *United States Elections Project*. Available from http://elections.gmu.edu/Turnout_2008G.html (accessed August 13, 2012).

McDonald, Michael P. (2013). 2012 General Election Turnout Rates. *United States Elections Project*. Available from http://elections.gmu.edu/Turnout_2012G.html (accessed March 8, 2013).

Silver, Nate (2012). FiveThirtyEight's 2012 forecast. http://fivethirtyeight.blogs.nytimes.com/fivethirtyeights-2012-forecast/ (accessed March 8, 2013).

StataCorp. (2009a). *Stata base reference manual: Release 11*. College Station, TX: StataCorp.

StataCorp. (2009b). *Stata survey data reference manual: Release 11*. College Station, TX: StataCorp.

Table 1: Distribution of subjective probabilities by selected survey weeks (%)

Survey week	Subj.prob. of voting			Subj.prob. of voting for Obama, conditional on voting		
	0%	1%-99%	100%	0%	1%-99%	100%
07/25-07/31	3.2	34.3	62.5	33.0	36.2	30.8
08/22-08/28	4.5	30.3	65.2	34.6	33.6	31.7
09/19-09/25	3.4	28.5	68.1	32.0	32.2	35.9
10/24-10/30	4.2	20.2	75.6	37.8	22.3	39.9

Table 2: Relationship between subjective and actual probabilities

	Survey week			
	07/25-07/31	08/22-08/28	09/19-09/25	10/24-10/30
Probability of voting				
Subj.prob. of voting	0.850***	0.893***	0.938***	1.001***
	(0.028)	(0.024)	(0.027)	(0.018)
Constant	12.626***	8.663***	3.494	-2.859
	(2.748)	(2.441)	(2.679)	(1.767)
R-squared	0.44	0.53	0.51	0.63
N	2723	2599	2195	2090
Probability of voting Obama, conditional on voting				
Without additional controls				
Subj.prob. of voting Obama	0.969***	0.984***	0.991***	1.005***
(conditional on voting)	(0.008)	(0.007)	(0.007)	(0.006)
Constant	4.304***	3.823***	1.383*	0.68
	(0.727)	(0.696)	(0.712)	(0.598)
R-squared	0.75	0.78	0.80	0.85
N	2411	2293	1935	1798
With subj. prob of voting				
Subj.prob. of voting Obama	0.971***	0.984***	0.991***	1.004***
(conditional on voting)	(0.008)	(0.007)	(0.007)	(0.006)
Subj.prob. of voting	-0.237***	-0.160***	-0.210***	-0.167*
	(0.052)	(0.061)	(0.079)	(0.090)
Constant	26.782***	19.183***	21.683***	17.014*
	(5.117)	(6.087)	(7.759)	(8.953)

R-squared	0.75	0.78	0.80	0.86
N	2411	2291	1933	1798

* p<0.10, ** p<0.05, *** p<0.01

Figure 1: Cumulative participation and sample freeze

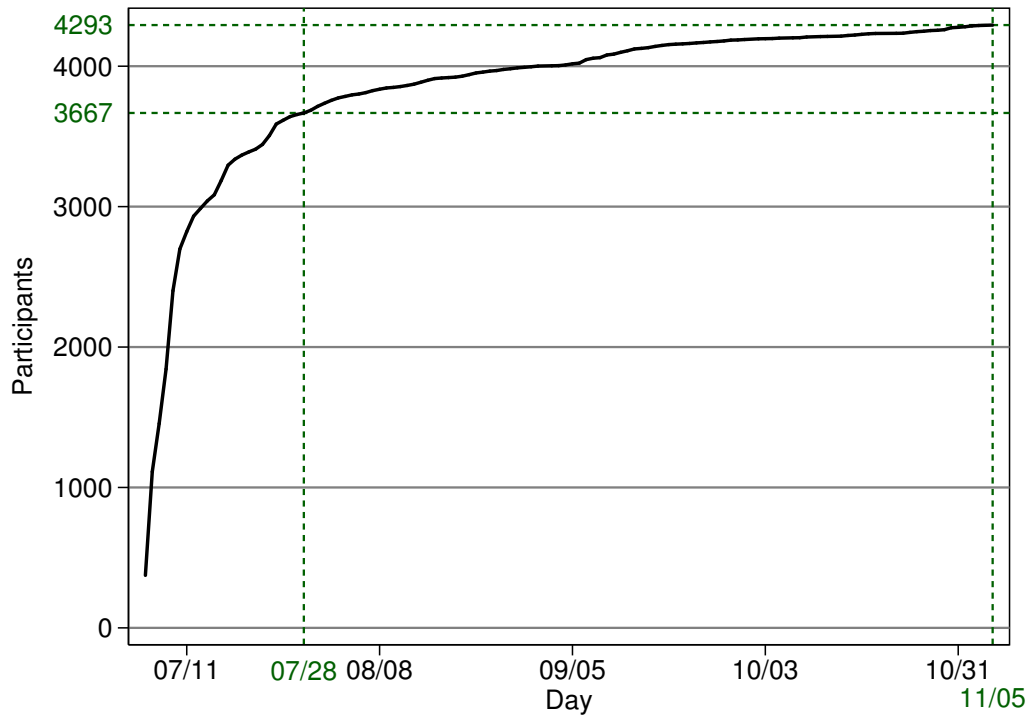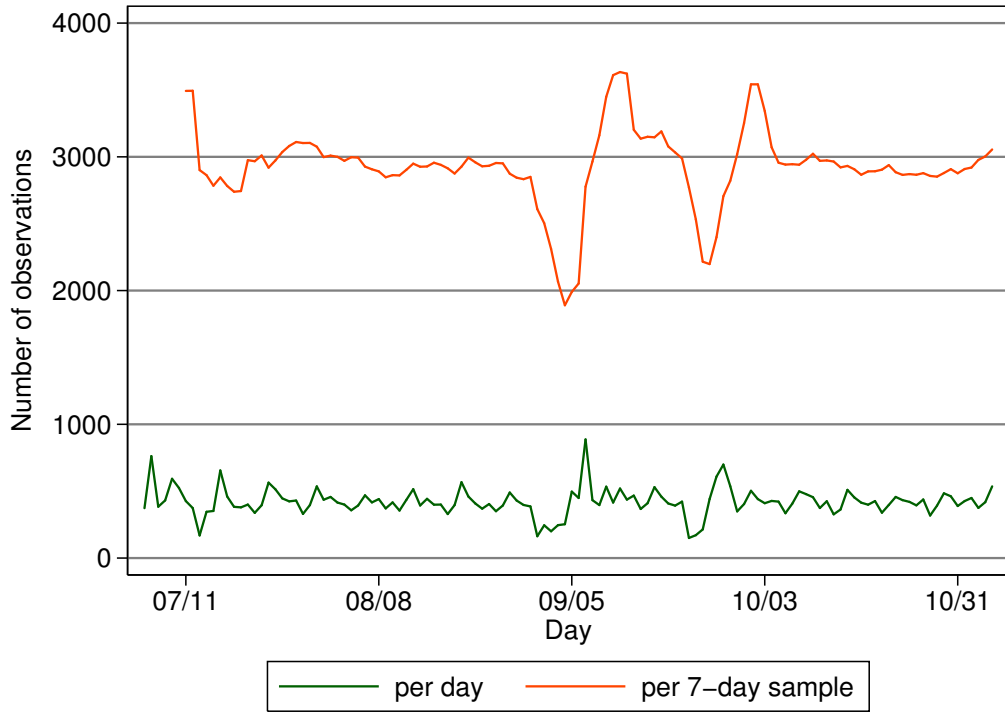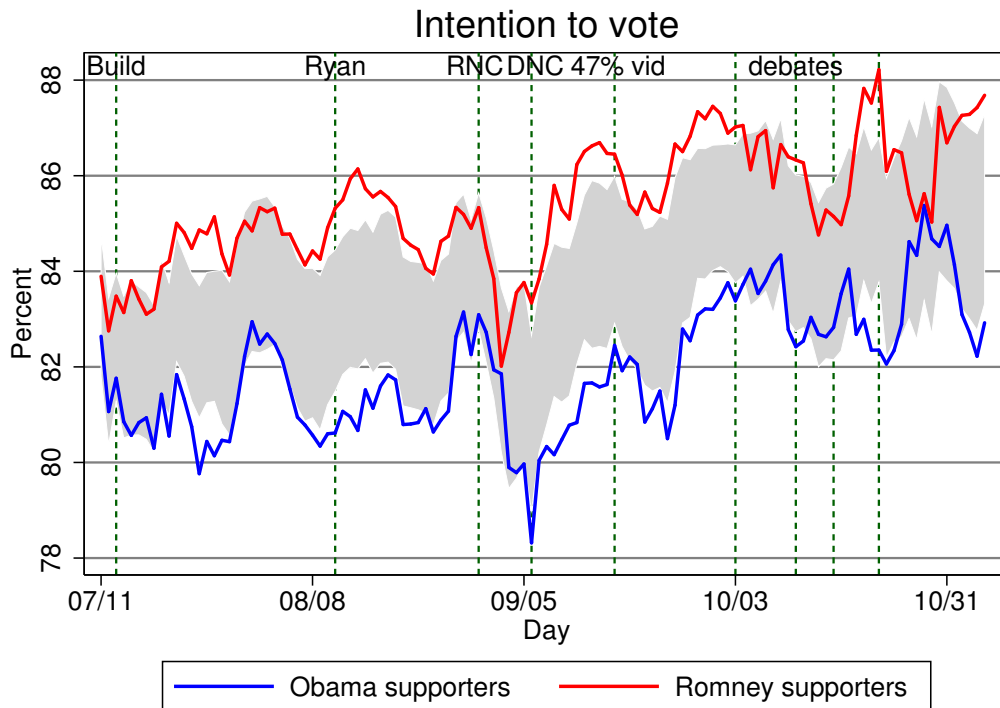**Figure 2:** Daily sample sizes and sample sizes for seven-day samples

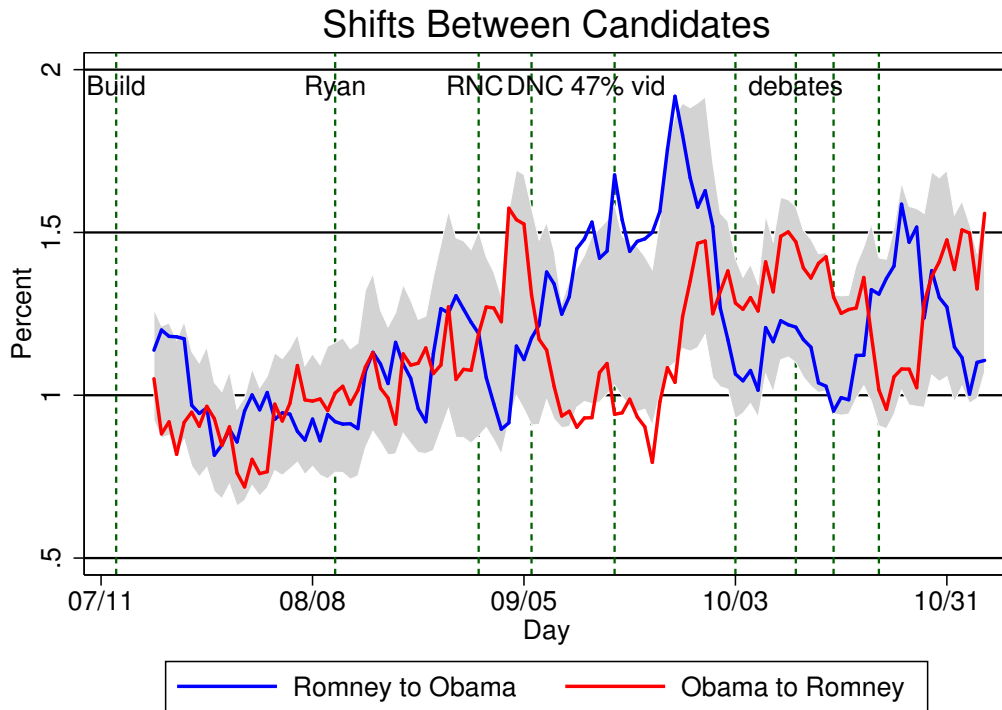**Figure 3:** Average stated intention to vote by candidate preference
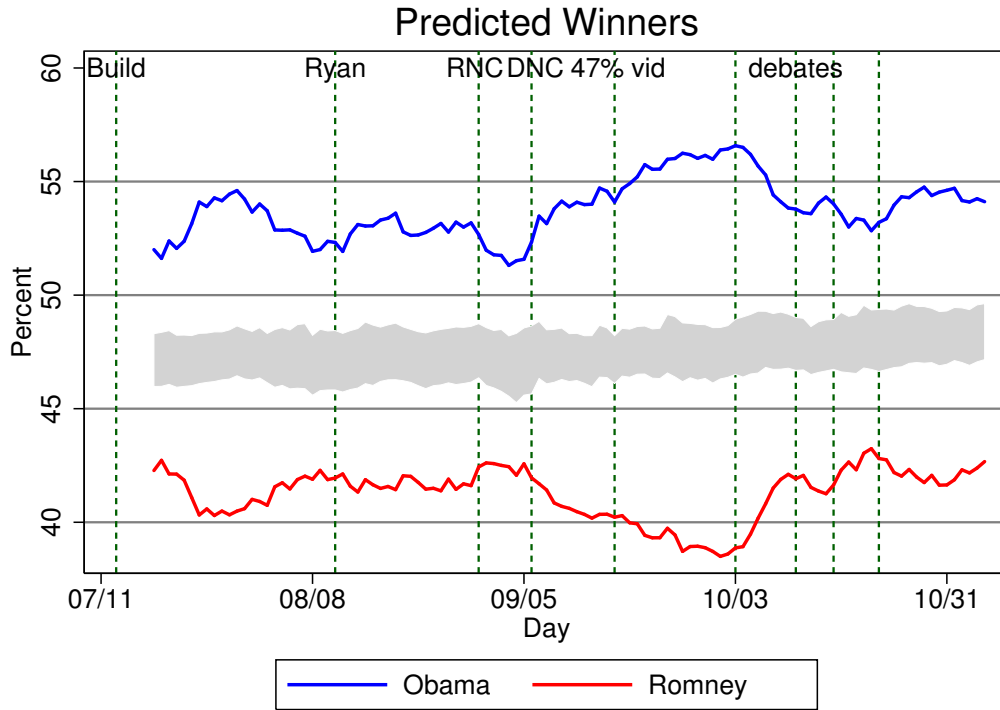

Note. An estimate within the shaded area indicates that the difference is not statistically significant at the 5% level. See text for a brief description of the indicated events.

Figure 4: Forecasted share of the votes for Obama and Romney
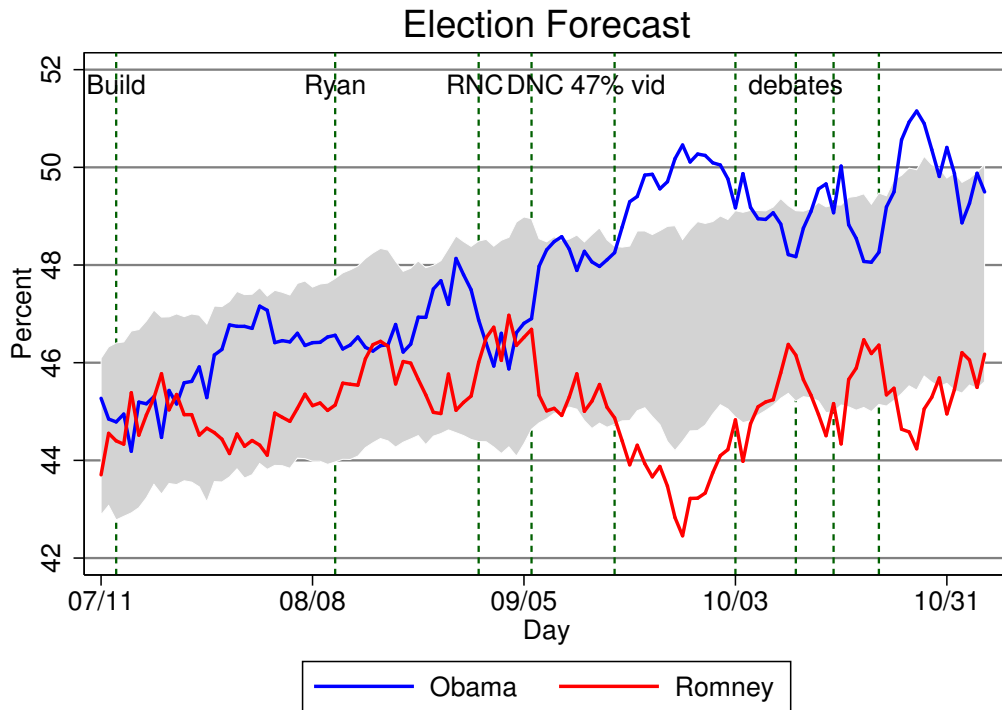
Note. An estimate within the shaded area indicates that the difference is not statistically significant at the 5% level. See

text for a brief description of the indicated events.

Figure 5: Average changes in reported likelihood of voting for Obama or Romney

Note. An estimate within the shaded area indicates that the difference is not statistically significant at the 5% level. See text for a brief description of the indicated events.

Figure 6: Average subjective likelihoods that Obama or Romney will win the election

Note. An estimate within the shaded area indicates that the difference is not statistically significant at the 5% level. See text for a brief description of the indicated events.

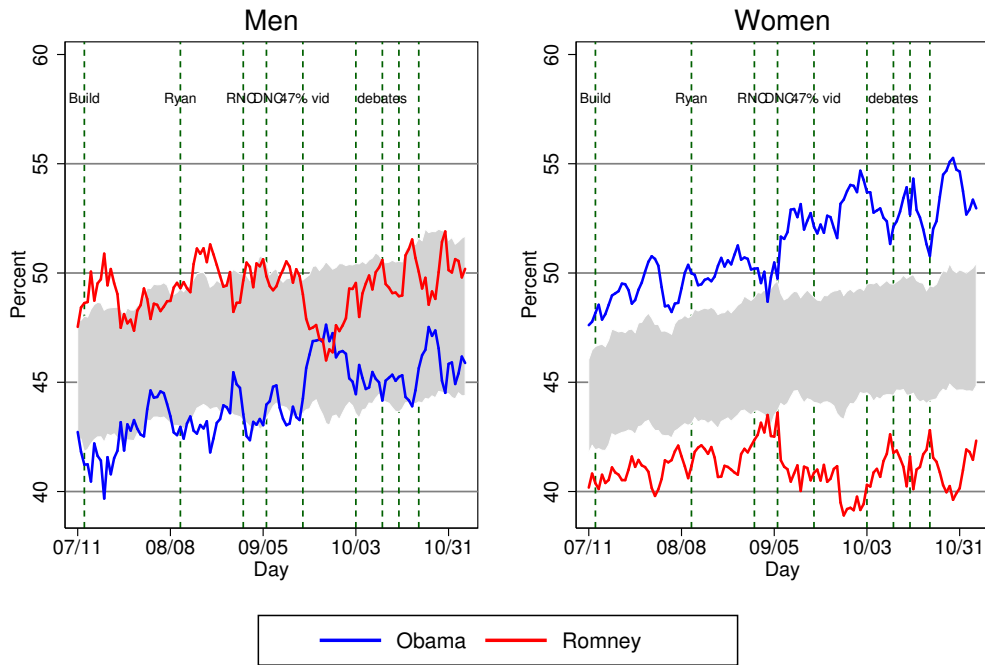**Figure 8:** Forecasted share of the votes for Obama and Romney, by age