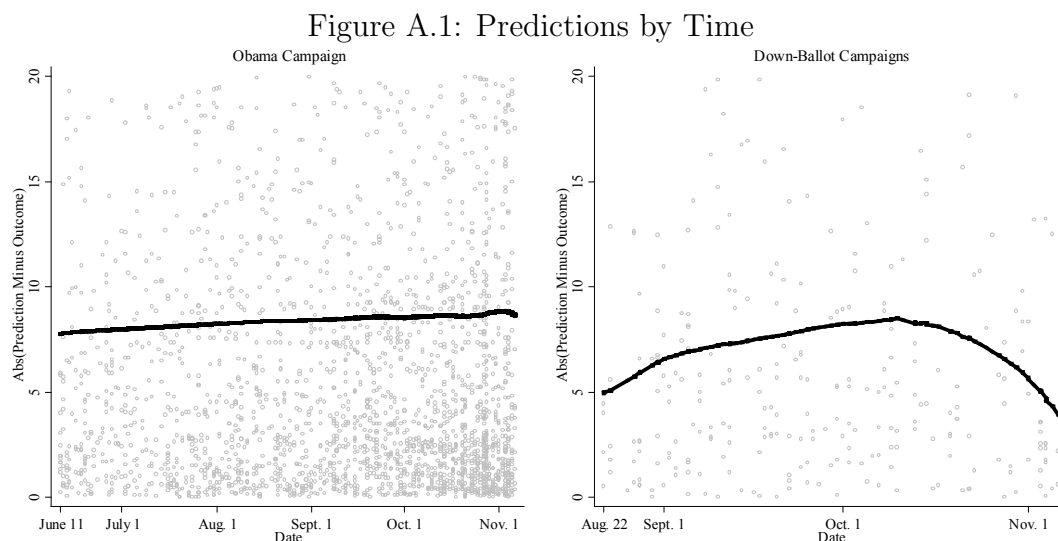


# A Supporting Information

In this appendix, we first provide supplementary information discussed in the body of the paper. Then, we provide additional information about the survey design and the sample.

## A.1 Additional Analysis



Note: Note that the x-axes in the two plots are scaled differently because we began interviewing Obama workers in June and down-ballot workers in August.

In discussing the results in Table 1, we mention a general absence of linear time effects. In other words, though we interviewed campaign workers for months preceding the election, it was not generally the case that predictions improved over time. Figure A.1 shows the difference between predictions and outcomes for individual respondents over time. Note that for clearer display, the figure trims predictions that are off by more than 20 points, but the lowest curve incorporates all responses. In the Presidential campaign, there is essentially no time trend. There is a slight increase in error as time goes on. In the down-ballot campaigns there is a curvilinear relationship with time. The best predictions are in early August and early November, with worse predictions in September and October. This same pattern is visible in the regression table (Table 1), even when controlling for other variables.

## A.2 Survey Design and Response Rates

Working with the Obama campaign and NGP-VAN, it was determined that the best way to survey workers was through a part of the NGP-VAN website called the Grid View. This is the place on the website where workers look up voters and can enter information about voters. For example, if a campaign worker calls a voter and the voter says they were supporting President Obama or that they would like to volunteer, this information could be entered into the database in the Grid View. For the Obama campaign, we solicited workers with a sampling probability of 1 in 100. For every log-in to the Grid View, a worker had a 1 in a 100 chance of being surveyed. A pop-up screen showed up on these workers' computers. If a user clicked "Okay," a web browser opened, and they viewed an informed consent page, followed by the five-minute survey. If the user clicked "Ask me later," they retained a 1 in 100 probability of being solicited for the survey. The survey went into the field for Obama staffers on June 11, 2012, and staffers were surveyed through November 6, 2012.

Working with the Association of State Democratic Chairs, we gained access to down-ballot races in 25 of the 49 states that use NGP-VAN. The included states are Alabama, Alaska, Arizona, Colorado, Connecticut, Idaho, Iowa, Indiana, Kansas, Kentucky, Louisiana, Massachusetts, Michigan, Mississippi, Nebraska, Nevada, New Mexico, North Carolina, Ohio, Oregon, South Dakota, Virginia, Vermont, Washington and Wyoming. Because the volume of use in down-ballot races is lower than in the Obama campaign (i.e. some campaigns may log-in to NGP-VAN only periodically), our sampling probability for down-ballot races was 1 in 33 rather than 1 in 100.

Between June 11th and November 6th, we solicited 15,953 distinct individuals. Of these, 5,608 (35.2%) entered the survey. While 5,608 clicked through to the survey, not all participated in every question and some seem to have clicked through to the survey and closed the window at the informed consent page. For most survey questions, our sample size is in the range of 3,000-3,500 resulting in an item-level response rate of approximate 20%. Not all respondents could be identified with a single campaign. For example, workers associated with state parties and with coordinated efforts might have worked on multiple campaigns. These respondents are excluded from most of our analysis in this manuscript because they had no single race for which they could predict an outcome.

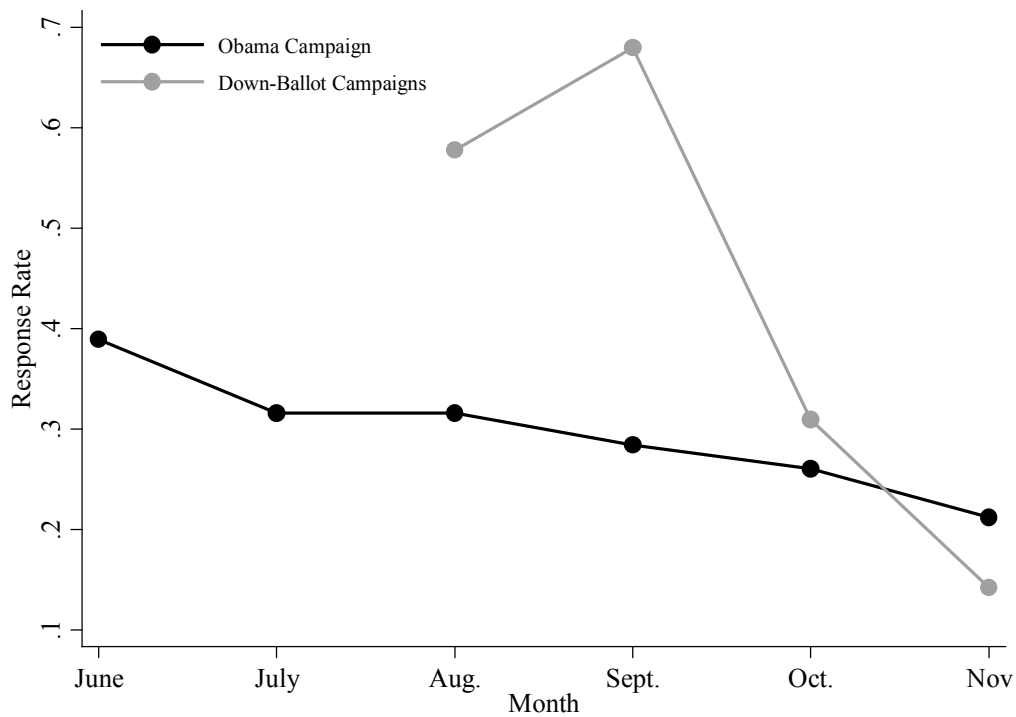
Table A.1 and Figure A.2 show two views at our response rate. In Table A.1, we show the rate of click-throughs to our survey by the number of solicitations. Recall that if a user clicked "Ask Me Later" at the initial survey invitation, they were still asked to take the survey in the future with the

Table A.1: Response Rate by Number of Solicitations

Solicitations	Response Rate	Observations
One	40.8	10,642
Two	26.2	2,645
Three	22.1	1,152
Four+	20.8	1,514

Note: Of those solicited four or more times, the median was solicited five times.

Figure A.2: Response Rate by Month of Interview



Note: Counts of solicitations by month for the Obama campaign are June (901), July (1,404), August (1,643), September (2,430), October (3,845) and November (1,014). Counts of solicitations by month for down-ballot campaigns are August (787), September (2,188), October (881), and November (647).

same initial sampling probability. As Table A.1 shows, over 40% of workers clicked through to the survey the first time they saw the prompt. Two-thirds of the solicitations were only solicited once. Individuals who refused the first

solicitation tended to refuse subsequent solicitations, as is shown in the table. However, we did capture a number of respondents who refused multiple times before agreeing to participate.

Figure A.2 shows the response rate over time. For both Obama workers and down-ballot workers, response rates were higher earlier in the campaign season than later. The behavior in down-ballot races is clearly very different than the Obama race. In the Obama campaign, the solicitations increased over time. Therefore, while there is a modest decline in response rate, our sample grows over time because the population expands. In down-ballot races, the pool of workers is largest in September and November, and smaller in October. (Remember that the November pool only includes six days of the month.) From talking with campaign strategists, the down-ballot pattern is likely attributable to the fact that down-ballot campaigns are run in a more staggered manner than the Presidential campaign. Campaigns may draw lists of voters and volunteers in September and then work off of those lists for the remainder of the campaign. Whereas at the Presidential level, with its greater resources, there is a continuous updating of voter lists and information. This may explain the pattern shown in Figure A.2.

Table A.2: Respondents by Race Type

Race Type	Observations
Obama Campaign	2,884
U.S. Senate/House	303
State House	156
Other State/Local and Coordinated	125

Table A.3: Respondents by Campaign Role

Role	Obama Campaign	Down-Ballot Campaign
Staff	454	228
Intern	269	48
Volunteer	2108	222
Candidate		15
Spouse		2
Other	129	15

Finally, in Tables A.2, A.3, and A.4, we show some key summary statistics from the sample. In particular we show the composition of the sample

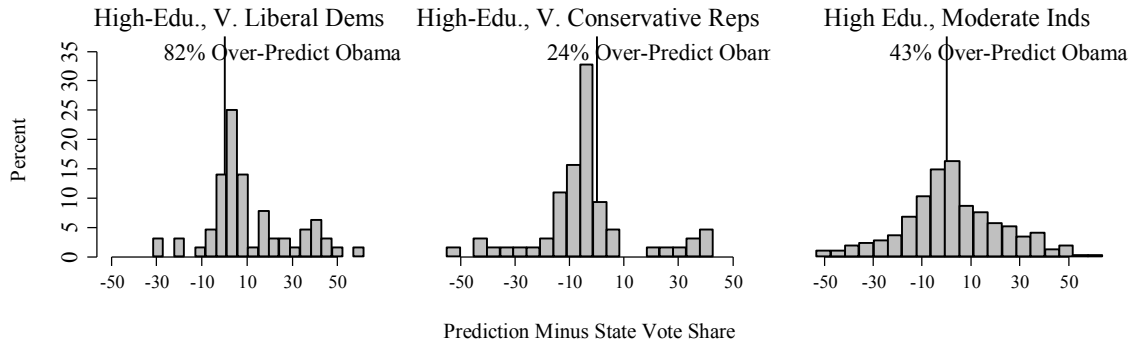
Table A.4: Staffer Respondents by Title

Title	Obama Campaign	Down-Ballot Campaign
Campaign Manager	0	39
Field Organizer	347	108
Regional Field Director	16	21
Deputy Field Organizer	64	7
Data Staff	5	8
Other	21	45

by race type (e.g. Presidential, federal, state), by role on the campaign (e.g. staff, intern, volunteer), and by staff title (e.g. campaign manger, field organizer). The key takeaway from these table is that while the Obama campaign respondents are overwhelmingly composed of volunteers, most of the down-ballot respondents are actually staffers. And of these, a substantial portion are higher level staffers. Even so, as referenced in the text, higher level and lower level staffers are about as good as one another at predicting the election outcome in their race.

### A.3 Replication of Figure 7 with High-Education Ideologues

Figure A.3: Partisanship Drives Responses Even Among Highly-Educated Partisans



Note: Data come from a combined sample of two 1,000 person modules of the 2012 CCES and 2012 CCAP. Observation counts for this survey item are 64 very liberal, college-graduate Democrats, 64 very conservative college-graduate Republicans, and 90 moderate, college-graduate Independents. Vertical line at  $x = 0$  indicates perfect prediction.

## A.4 Replication of Table 1 with Variable for First Campaign

Table A.5: Which Campaign Workers Accurately Predict Election Outcomes?

Ind. Variables	Obama Campaign			Down-Ballot Campaigns		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
College	-3.80*** (0.91)	-4.07*** (0.93)	-5.27*** (1.06)	-4.66*** (1.75)	-3.88** (1.94)	-5.09** (1.96)
Paid Staff	-2.31*** (0.76)	-1.78** (0.71)	-1.28* (0.73)	1.13 (1.11)	0.56 (1.20)	-1.54 (1.62)
First Campaign	1.10** (0.43)	1.20*** (0.40)	1.23*** (0.44)	1.17 (1.20)	1.10 (1.17)	0.47 (1.36)
V. Liberal	-0.36 (0.40)	-0.3 (0.41)	-0.46 (0.42)	-2.17** (0.99)	-1.81* (0.97)	-1.52 (1.28)
Num. Polls in State		-0.08*** (0.02)	-0.09*** (0.02)		-0.06** (0.03)	
$\Delta$ Obama St. Adv.			0.05 (0.16)			
Incumbent					-0.1 (2.22)	0.45 (1.74)
Federal Race					-2.15 (1.50)	-4.27** (1.70)
Summer			0.81 (0.94)		2.41 (2.08)	2.27 (2.60)
September			2.52*** (0.86)		1.24 (1.68)	0.63 (2.34)
October			2.24** (0.97)		3.99** (1.58)	3.21 (2.25)
Intercept	11.51*** (0.79)	15.80*** (1.41)	16.03*** (2.04)	11.34*** (1.72)	11.60*** (2.06)	13.59*** (3.06)
Obs.	1,564	1,564	1,257	137	136	136
$R^2$	0.03	0.07	0.09	0.09	0.18	0.36

Note: dependent variable is  $|\text{prediction} - \text{outcome}|$ . Standard errors are in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Models 1-3 employ robust clustered standard errors, where the state of Obama workers is the clustering variable. Models 4-5 cluster standard errors around the specific down-ballot campaign of the worker. For an alternative specification, Model 6 employs state fixed effects instead of clustered standard errors. The number of presidential polls in the state is collinear with the set of state indicators in Model 6.