Supplementary Information to

Is Voting Habit Forming? New Evidence from Experiments and Regression Discontinuities

Alexander Coppock and Donald P. Green

April 28, 2015

This online appendix to *Is Voting Habit Forming? New Evidence from Experiments and Regression Discontinuities* comprises five sections: a discussion of challenges associated with using voting-age eligibility discontinuities to identify the habit-forming effects of voting with particular attention paid to biases resulting from measurement error, an exploration of the robustness of our results to alternative specifications, and an extension of the discontinuity analysis to primary elections, a presentation of ANES survey results showing the null effects of eligibility on campaign contact, and a further description of the follow-up experiment conducted with the subjects of the 2007 social pressure experiment.

1 Identification of the CACE

Our point of departure is the potential outcomes framework that has been used to investigate identification and estimation in the context of experiments that confront two-sided noncompliance (Angrist, Imbens and Rubin 1996). Potential outcomes are fixed attributes of each individual that indicate what he or she would do if exposed to a particular treatment or combination of treatments. Before discussing the downstream effect of voting on subsequent turnout, let us first consider the relationship between encouragement (or eligibility) to vote and turnout in the first election. Suppose that each person i harbors two potential outcomes that indicate whether he or she would vote in the upstream election if exposed to an encouragement ($Z = 1$) or not ($Z = 0$). Let $V_{1i}(1)$ be i's turnout in the upstream election if i is exposed to the encouragement, and $V_{1i}(0)$ be i's turnout in the upstream election if i is not exposed to the encouragement.[1]

The treatment effect of the encouragement on turnout in the upstream election is defined as:

$$t_i = V_{1i}(1) - V_{1i}(0) \tag{A1}$$

In other words, the causal effect is defined as the difference between two potential states of the world, one in which the individual receives the encouragement, and another in which the individual does not. Extending this logic from a single individual to a set of individuals, we define the average upstream treatment effect (AUTE) of the encouragement as follows:

$$AUTE = \mathrm{E}\left[t_i\right] = \mathrm{E}\left[V_{1i}(1) - V_{1i}(0)\right] \tag{A2}$$

where $\mathrm{E}\left[\cdot\right]$ indicates an expectation over all subjects.

In an actual experiment, we observe subjects either receiving an encouragement or not, never both. The causal effect for a given individual, expressed in equation A1, cannot be estimated, but random assignment of the encouragement allows us to obtain unbiased estimates of the AUTE in equation A2. Let Z_i denote whether an individual is encouraged to vote. The difference in expected outcomes among those who are encouraged and those who are not may be expressed as:

$$\mathrm{E}\left[V_{1i}(1)|Z_i = 1\right] - \mathrm{E}\left[V_{1i}(0)|Z_i = 0\right] \tag{A3}$$

where the notation $\mathrm{E}\left[A_i|B_i = c\right]$ means the average value of A_i among those subjects for which

[1]When characterizing potential outcomes in this way, we implicitly invoke the stable unit treatment value assumption (Rubin 1990), which stipulates that potential outcomes do not depend on which subjects are assigned to treatment. This assumption would be jeopardized, for example, when the treatment administered to one subject affects the outcomes of other subjects.

the condition $B_i = c$ holds. When random assignment determines which treatment each subject receives, Z_i is independent of potential outcomes. The expected difference-in-means is therefore equal to the average treatment effect:

$$\mathrm{E}\left[V_{1i}(1)|Z_i = 1\right] - \mathrm{E}\left[V_{1i}(0)|Z_i = 0\right] = \mathrm{E}\left[V_{1i}(1) - V_{1i}(0)\right] = AUTE \tag{A4}$$

This result demonstrates the fact that randomized experiments generate unbiased estimates of the effects of encouragement on voting in the first election. However, the focus here is not on the proximal effect of encouragement; rather, we aim to estimate the causal effect of voting in the upstream election on voting in a subsequent election. Let $V_{2i}(Z, V_1)$ denote subject i's potential vote or abstention in the second election. This potential outcome responds to two inputs: whether the subject received an encouragement prior to the first election and whether the subject voted in the first election. As Angrist, Imbens and Rubin (1996) point out, our ability to identify the causal effect of voting in the first election hinges on a crucial simplifying assumption, known as an exclusion restriction. The assumption states that $V_{2i}(Z, V_1) = V_{2i}(V_1)$, which is to say that subjects' turnout in the second election responds to whether they vote in the first election, but not to whether they are encouraged to vote prior to the first election. This assumption would be violated, for example, if subjects' vote in the second election were influenced by whether they received encouragement, holding constant whether they voted in the first election. As noted earlier, this assumption may be jeopardized, for example, when memorable social pressure messages are used to encourage turnout.

In order to recover the causal effect of V_{1i} on V_{2i} we need one further assumption known as monotonicity (Angrist, Imbens and Rubin 1996). Describing this assumption requires a bit more terminology. Depending on the way their votes in the upstream election potentially respond to encouragement, subjects may be classified into four types, compliers, never-takers, always-takers, and defiers. Compliers are subjects who vote in the upstream election if and only if assigned to the encouragement. For this group $V_{1i}(1) - V_{1i}(0) = 1$. Never-takers are those who would not vote in the upstream election no matter their assignment: $V_{1i}(1) = V_{1i}(0) = 0$. Conversely, always-takers are those who vote in the upstream election no matter their assignment: $V_{1i}(1) = V_{1i}(0) = 1$. Defiers are those who vote in the upstream election if and only if they are assigned to the control group: $V_{1i}(1) - V_{1i}(0) = -1$. The monotonicity assumption stipulates that there are no defiers. In context of habit research, monotonicity holds that the encouragement to vote either has no upstream effect or causes a person to vote in an upstream election who would otherwise abstain.

Under the exclusion restriction and monotonicity assumptions, a randomized experiment that generates some positive share of compliers (i.e., there must be some effect of encouragement on V_{1i}) identifies the ATE among compliers (Angrist, Imbens and Rubin 1996). This quantity, the Complier Average Causal Effect (CACE), refers to the average effect of voting in the upstream election on voting in the downstream election among a subset, those who would vote upstream if and only if encouraged to do so. The effect encompasses the entire causal process that is set in motion by the upstream election – including voting in intermediate elections. We are only able to estimate the total effect of the upstream election on downstream behavior; we cannot attribute portions of the effect to the intermediate elections. The CACE is estimated by dividing two sample estimates. The numerator in equation A5 is the average downstream vote in the assigned treatment group minus the average downstream vote in the assigned control group; the denominator is the observed upstream vote in the assigned treatment group minus the observed upstream vote in the control group:

$$\widehat{CACE} = \frac{\hat{\mathrm{E}}\left[V_{2i}|Z_i = 1\right] - \hat{\mathrm{E}}\left[V_{2i}|Z_i = 0\right]}{\hat{\mathrm{E}}\left[V_{1i}|Z_i = 1\right] - \hat{\mathrm{E}}\left[V_{1i}|Z_i = 0\right]} \tag{A5}$$

This ratio is equivalent to the estimate generated by an instrumental variables regression of V_{2i} on V_{1i} using Z_i as an instrumental variable. Because the denominator is a difference between two quantities that are subject to sampling variability, this ratio is consistent but not unbiased and becomes undefined when the encouraged group and the non-encouraged group have the same voting rates in the first election. Precise estimation therefore requires that the encouragements bear a reasonably strong relationship to voting in the first election. Indeed, the stronger the relationship between encouragement and voting in the first election, the more resilient the estimates are to biases associated with violations of the exclusion restriction (Conley, Hansen and Rossi 2012). There is little to be learned from investigating the downstream effects of weak manipulations, but we note the potential for selection bias were one to analyze only experiments with statistically significant first-stage relationships. The experimental results presented here sidestep this problem, first because the very large N ensures no appreciable sampling error, and second because we present all of our replications of this study regardless of outcomes.

This framework may also be applied to research designs in which a discontinuity is used to identify the habit effect. In place of random assignment, we assume that those just over the legal voting age have the same expected potential outcomes as those just under the legal voting age. The analogy between discontinuities and random assignments is plausible in this application, as subjects or administrators are unlikely to take actions to alter assignments (e.g., by falsifying birthdates so that certain subjects can vote). The use of birthdate cutoffs also satisfies the monotonicity assumption insofar as the assigned control group (i.e., those ineligible to vote in the first election) cannot vote, which implies that there can be no defiers. The crucial statistical assumption is again the exclusion restriction, which holds that the only way that eligibility to vote in the first election influences voting in the second election is via the act of voting in the first election, not some backdoor path that stems from voting-age eligibility.

Since both experimental designs and discontinuities identify CACEs, in principle estimation proceeds the same way in discontinuity analysis as it does in experimental analysis: consistent estimates are obtained via a 2SLS regression of voting in the second election on voting in the first election, using the assigned treatment (eligibility to vote in the first election) as an instrument. However, as Meredith (2009) points out, the manner in which voting records are assembled in the US introduces a complication. Because there exists no official list of 17-year-olds, we do not observe the number of subjects who fall just short of the age eligibility cutoff. Moreover, we do not have a comprehensive list of all just-eligible 18-year-olds; we have a list of those who registered to vote by some later date.

Fortunately, we possess the total number of votes cast by people who were just over or under the eligibility threshold. If we assume that the birthdate cutoff partitions the subject pool randomly, it follows that the number of subjects who are just above the cutoff will, in expectation, be the same as the number of subjects who are just below the cutoff. Aronow (2013) shows that the assumption that $Pr(Z_i = 1) = Pr(Z_i = 0)$ is sufficient to identify the CACE under the exclusion restriction, monotonicity, and random assignment. The estimator is similar to equation A5, except the numerator is the difference in total votes in the second election and the denominator is the difference in total votes cast in the first election. Using N_{18} to denote the number of just-eligible 18-year-olds and N_{17} to denote the just-ineligible number of 17-year-olds:

$$\widehat{CACE} = \frac{\hat{\mathrm{E}}\left[V_{2i}|Z_i = 1\right] - \hat{\mathrm{E}}\left[V_{2i}|Z_i = 0\right]}{\hat{\mathrm{E}}\left[V_{1i}|Z_i = 1\right] - \hat{\mathrm{E}}\left[V_{1i}|Z_i = 0\right]}$$

$$= \frac{\frac{\sum_{i=1}^{N_{18}} V_{2i}(Z_i = 1)}{N_{18}} - \frac{\sum_{1}^{N_{17}} V_{2i}(Z_i = 0)}{N_{17}}}{\frac{\sum_{i=1}^{N_{18}} V_{1i}(Z_i = 1)}{N_{18}} - \frac{\sum_{1}^{N_{17}} V_{1i}(Z_i = 0)}{N_{17}}}$$

Under the assumption that $N_{18} = N_{17} = N$:

$$= \frac{\frac{\sum_{i=1}^{N} V_{2i}(Z_i = 1)}{N} - \frac{\sum_{1}^{N} V_{2i}(Z_i = 0)}{N}}{\frac{\sum_{i=1}^{N} V_{1i}(Z_i = 1)}{N} - \frac{\sum_{1}^{N} V_{1i}(Z_i = 0)}{N}},$$

which reduces to

$$\widehat{CACE} = \frac{\sum_{1}^{N} V_{2i}(Z_i = 1) - \sum_{1}^{N} V_{2i}(Z_i = 0)}{\sum_{1}^{N} V_{1i}(Z_i = 1) - \sum_{1}^{N} V_{1i}(Z_i = 0)}. \tag{A6}$$

Notice that this estimator does not require information about the number of people falling just above or below the eligibility cutoff. So long as we assume that this number is the same just above and below the cutoff, this constant cancels in both the numerator and the denominator. Since the number of subjects whose birthdays fall on either side of the cutoff may be small, regression with linear time trends is used to make plausible the assumption that the just-eligible and just-ineligible have potential outcomes that are identical in expectation. The unit of analysis is the group of people who are born on a given day. In order to eliminate seasonal trends, we include the number of votes cast by the birth cohort born on the same day of the week approximately one year earlier as a regressor.[2] The regression discontinuity model comprises two equations:

$$\text{Downstream Votes Cast} = \beta_0 + \beta_1 Z_j + \beta_2 T_j + \beta_3 Z_j * T_j + \beta_4 \text{Lagged Downstream}_j + \epsilon_j \quad \text{(A7)}$$

$$\text{Upstream Votes Cast} = \alpha_0 + \alpha_1 Z_j + \alpha_2 T_j + \alpha_3 Z_j * T_j + \alpha_4 \text{Lagged Downstream}_j + \eta_j \quad \text{(A8)}$$

where Z_j is an indicator for eligibility to vote in the upstream election and T_j is a running variable indicating the number of days between a birthday and the eligibility cutoff. An estimate of the CACE can be obtained by estimating both equations by OLS and calculating the ratio of β_1 to α_1:

$$\widehat{CACE} = \frac{\hat{\beta_1}}{\hat{\alpha_1}} \quad \text{(A9)}$$

More directly, we can estimate the system of equations implied by equations A7 and A8 using two-stage least squares, which has the advantage of allowing us to easily estimate a robust standard error.

1.1 Measurement Error: Implications for Estimation of the CACE

One potential challenge faced by our analytic strategy is the measurement error caused by residential mobility. There are two ways for data to be censored from the voter file. The first is the "deadwood" problem in which a voter remains on the file when she has moved to another state. The second is when a voter moves into the state, but her voter history does not travel with her. Having lost her vote history, she is marked as not having voted in any elections prior to the move, which may or may not be correct. This mismeasurement can lead to biased estimates of the Average Upstream Treatment Effect (AUTE) or the Average Downstream Treatment Effect (ADTE), thereby biasing our estimates of the Complier Average Causal Effect (CACE). A voter

[2] Correctly coding the prior year's birthdate is complicated by leap days, which are excluded from the analysis.

can move before both the upstream and downstream elections occur, in between the two elections, or after both. The first scenario poses no problem, because her voter history will be accurately recorded in her new state of residence. The second two scenarios have different implications for bias, and we will consider them in turn.

1.1.1 Case 1: Voter Moves in between Upstream and Downstream Elections

The inferential target is the CACE among those who reside in the state at the time the voter file was produced. Therefore, anyone who voted in the upstream election out-of-state and then moved into the state would cause us to overestimate the CACE. By the same token, however, anyone who votes in-state in the upstream election and then moves out-of-state causes us to underestimate the CACE. The proportion of nevertakers who change states does not affect our estimates, though the proportion of compliers who move does.

Table A1 shows the 12 possible types of voters according to their migration status and voting behavior. For example, the first row shows a voter who a) is treated, b) votes in the upstream election out of state and c) votes in state in the downstream election. Incomplete data on the voter file leads us to mis-code this individual's upstream decision as a "0" but correctly mark the downstream vote as a "1". Had this person been in the control condition (and therefore in the 5th row of the table), there would be no problem, because control subjects cannot vote upstream regardless of their state of residence. The table shows that the only subjects who cause us difficulty are treated compliers who move.

One subtle feature of our enumeration of types is that we do not include residents who move out of state between the upstream and downstream elections and whose upstream voting records are *appropriately* purged by the secretary of state. Because we define our estimand as the CACE among residents in a state at the time the voter file was generated, these individuals are correctly excluded from the estimation.

Table A1: Biases due to Residentially Mobile Voters

	Type	Z	Upstream Vote: True	Upstream Vote: Observed	Downstream Vote: True	Downstream Vote: Observed	Direction of Bias
In-migrant	1	1	1	0	1	1	Positive
	2	1	1	0	0	0	Positive
	3	1	0	0	1	1	None
	4	1	0	0	0	0	None
	5	0	0	0	1	1	None
	6	0	0	0	0	0	None
Out-migrant	7	1	1	1	1	0	Negative
	8	1	1	1	0	0	Negative
	9	1	0	0	1	0	None
	10	1	0	0	0	0	None
	11	0	0	0	1	0	None
	12	0	0	0	0	0	None

Consider the following numerical example:

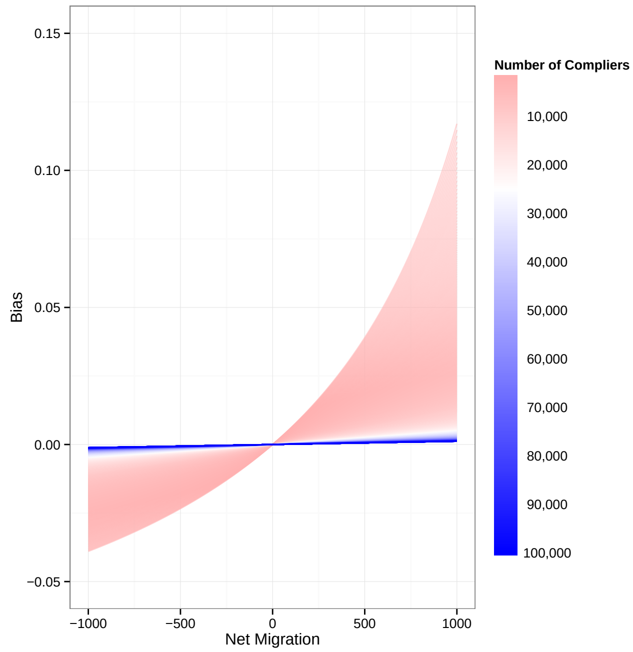
- The number of treated subjects who vote in the upstream election is 500. This is the number of compliers, written N_c.
- The number of treated subjects who vote in the downstream election is 1000.
- The number of untreated subjects who vote in the downstream election is 900.
- The true CACE is therefore $\frac{1000-900}{500} = 0.20$

Suppose that 5 treated compliers move in state between the upstream and downstream elections ($N_{mig} = 5$). We incorrectly estimate N_c to be 495. We therefore overestimate the CACE to be $\frac{1000-900}{495} = 0.202$.

Suppose now that 5 treated compliers move out of state between the upstream and downstream elections ($N_{mig} = 5$). We incorrectly estimate N_c to be 505. We therefore estimate the CACE to be $\frac{1000-900}{505} = 0.198$.

The magnitude of the bias due to migration depends on the true CACE, the number of compliers, and the net number of migrant compliers. The bias in each case can be calculated according to the following formula:

Figure A1: Bias due to Measurement Error (True CACE = 0.117)



$$Bias = \frac{CACE_{true} * N_{mig}}{N_c - N_{mig}} \tag{A10}$$

Figure A1 shows the extent of this bias under a wide range of circumstances. The true CACE in all cases is set equal to 0.117, the meta-analytic estimate of the CACE of 2008 voting on 2012 voting. We then vary the other two parameters in the bias formula: the number of compliers and the number of net migrants. The graph shows that as the ratio of migrant compliers to residentially stable compliers gets large, the bias becomes serious. States that experience net outflows are associated with negative biases; the opposite is true for states with net inflows.

In order to calibrate our understanding of where in this graph the majority of our estimates fall, we turn to a special tabulation obtained from Catalist in which voters are tracked across states. We use these data to construct a measure of the net migration flows in and out of each state for young people who voted in both the 2008 and 2012 elections. Table A2 presents the estimated number of net migrants, the estimated number of compliers in 2008 (computed using a 365-day window), and the associated bias according to Equation A10 for sixteen states. The table shows that the probable size of the bias due to measurement error is small, at least in the elections for which we have migration data.

Table A2: Migrants, Number of Compliers, and Associated Bias

State	Net Migrants	N_c 2008	CACE 08-12	Bias Estimate	Corrected CACE
PA	-1033	95175	0.121	-0.001	0.122
IA	-517	5318	0.086	-0.008	0.094
CT	-183	22413	0.161	-0.001	0.162
NV	-182	12465	0.174	-0.003	0.177
RI	-117	6867	0.113	-0.002	0.115
OK	-50	15584	0.138	-0.000	0.138
NJ	8	51983	0.155	0.000	0.155
MO	17	35535	0.155	0.000	0.155
MT	50	6770	0.111	0.001	0.110
KY	133	21964	0.075	0.000	0.075
OR	134	17853	0.108	0.001	0.107
AR	183	11778	0.200	0.003	0.197
NY	292	79908	0.068	0.000	0.068
FL	641	88519	0.105	0.001	0.104
IL	750	58991	0.080	0.001	0.079

1.1.2 Case 2: Voter Moves after both Upstream and Downstream Elections

The biases associated with residential mobility are different if, rather than moving in between the upstream and downstream elections, voters move after both elections have occurred. In-migrants are those who arrive in a state after both elections, but none of their vote history travels with them. These types are shown in the top panel of Table A3: regardless of their true voting behavior, they are scored as not having voted in either the upstream or downstream elections. Out-migrants, on the other hand, have their voting behavior measured correctly. The trouble is, however, they should be excluded from the estimation because our estimand is the CACE among those who reside in a state at the time the voter file was collected. Some out-migrants cause no bias because the secretaries of state correctly purge them from the voter lists. Other out-migrants remain on the file as deadwood and can lead to mismeasurement.

Table A3: Types of Voters Who Move After Upstream and Downstream Elections

	Type	Z	Upstream Vote True	Upstream Vote Observed	Downstream Vote True	Downstream Vote Observed	Direction of Bias
In-migrant	1	1	1	0	1	0	Negative
	2	1	1	0	0	0	Positive
	3	1	0	0	1	0	Negative
	4	1	0	0	0	0	None
	5	0	0	0	1	0	Positive
	6	0	0	0	0	0	None
Out-migrant	7	1	1	1	1	1	Positive
	8	1	1	1	0	0	Negative
	9	1	0	0	1	1	Positive
	10	1	0	0	0	0	None
	11	0	0	0	1	1	Negative
	12	0	0	0	0	0	None

To understand the probable magnitudes of the biases due to residential mobility, it is instructive to work through the implications of Table A3. We know from the as-if random assignment of Z, that the total number of types 1, 2, 3, and 4 should equal the total number of types 5 and 6. Similarly, the total number of 7's, 8's, 9's, and 10's should equal the 11's and 12's. Under the assumption that the habit effect is either zero or positive, we can make stronger statements:

$$N_1 + N_3 \geq N_5 \rightarrow \text{Net negative bias}$$

$$N_2 + N_4 \leq N_6 \rightarrow \text{Net positive bias}$$

$$N_7 + N_9 \geq N_{11} \rightarrow \text{Net positive bias}$$

$$N_6 + N_{10} \leq N_{12} \rightarrow \text{Net negative bias}$$

At this stage, we have to engage in some informed guesswork about the shares of the twelve types among those who move after both elections. First, we can be relatively sure that types 1-6 outnumber types 7-12 because most states do periodically clean their voter lists of nonresidents.

Second, because we are most concerned with the positive biases, a key question is how many type 2 voters there are. These are the just-eligible voters who vote in the upstream election but fail to vote in the downstream election. Given those who vote at age 18 tend to be unusually politically engaged, it seems safe to assume that relatively more of them would vote downstream than would not, i.e., the 1's outnumber the 2's by a fair amount.

We concede that measurement error may influence our CACE estimates, but we suspect that the magnitudes of these biases are likely to be small. We currently lack the data to make definitive statements about these biases. As data vendors become more adept at tracking voter histories across state lines, however, the requisite information may become available in the future.

1.2 Estimating Turnout Rates among Untreated Compliers

Tables 1, 2, and 3 include estimates of the voter turnout rate among untreated compliers, which is calculated according to the expression below (equivalent to Corollary 1 in Aronow and Green (2013, p.678)):

$$\frac{\mathrm{E}\left[V_{2i}(Z_i=0)\right]-\mathrm{E}\left[V_{1i}(Z_i=0)\right]*\mathrm{E}\left[V_{2i}(V_{1i}=1,Z_i=0)\right]-(1-\mathrm{E}\left[V_{1i}(Z_i=1)\right]*\mathrm{E}\left[V_{2i}(V_{1i}=0,Z_i=1]\right)}{\mathrm{E}\left[V_{1i}(Z_i=1)\right]-\mathrm{E}\left[V_{1i}(Z_i=0)\right]} \tag{A11}$$

These expectations may be estimated using sample analogues. Because estimates of turnout rate are subject to sampling variability, our estimates may fall outside of the possible range, in which case they were set to the boundary. This quantity may be estimated in a straightforward manner using experimental data. However, the regression discontinuity results do not include estimates of the turnout rate among untreated compliers because the voter files do not contain the required information, i.e., a full list of those who did not vote in the upstream and downstream elections.

1.3 Voter File Descriptive Statistics

Tables A4 and A5 show descriptive statistics for our voter files. Official turnout statistics were obtained from `http://www.electproject.org/`. Because voter files generally only include the voter history of those registered at the time the files were created, fewer votes are recorded for earlier elections. We obtained voter files for all states included in this study in 2013, but we included Florida and Missouri files obtained in 2005 for comparison. As voter files in the US become increasingly integrated, future work will be able to revisit the question of the habit-forming effects of voting with data less encumbered by measurement error concerns.

Table A4: Voter File Descriptive Statistics

Arizona 2013				Colorado 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992				1992	1,569,180	506,265	32%
1994				1994	1,116,307	410,794	37%
1996	884,262	800,567	91%	1996	1,510,704	720,474	48%
1998	706,011	654,788	93%	1998	1,327,235	866,317	65%
2000	921,781	906,092	98%	2000	1,741,368	1,199,980	69%
2002	805,696	796,836	99%	2002	1,416,093	1,125,752	79%
2004	1,054,945	1,052,094	100%	2004	2,129,630	1,867,515	88%
2006	774,680	767,034	99%	2006	1,558,405	1,498,031	96%
2008	1,086,617	1,079,987	99%	2008	2,401,361	2,407,910	100%
2010	781,333	784,631	100%	2010			
2012	1,069,468	1,069,824	100%	2012			

Connecticut 2013				Iowa 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992	1,616,332	21,826	1%	1992			
1994	1,147,084	63,064	5%	1994			
1996	1,392,614	125,741	9%	1996			
1998	999,537	126,590	13%	1998			
2000	1,459,525	414,242	28%	2000	1,315,563	674,740	51%
2002	1,022,942	395,885	39%	2002	1,023,075	783,598	77%
2004	1,578,769	973,078	62%	2004	1,506,908	897,129	60%
2006	1,134,780	809,714	71%	2006	1,048,033	901,715	86%
2008	1,646,792	1,521,693	92%	2008	1,537,123	1,342,435	87%
2010	1,153,115	1,114,094	97%	2010	1,116,063	1,059,694	95%
2012	1,558,960	1,510,273	97%	2012	1,582,180	1,548,066	98%

Florida 2005				Florida 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992	5,314,392	3,406,554	64%	1992	5,314,392	1,436,265	27%
1994	4,206,659	3,041,209	72%	1994	4,206,659	1,339,961	32%
1996	5,300,927	4,571,420	86%	1996	5,300,927	2,232,143	42%
1998	3,965,751	3,606,169	91%	1998	3,965,751	2,038,740	51%
2000	5,963,110	5,537,473	93%	2000	5,963,110	3,394,513	57%
2002	5,100,581	4,532,080	89%	2002	5,100,581	3,197,738	63%
2004	7,609,810	7,368,460	97%	2004	7,609,810	5,086,860	67%
2006				2006	4,829,270	5,406,450	112%
2008				2008	8,390,744	6,823,216	81%
2010				2010	5,411,106	4,761,288	88%
2012				2012	8,474,179	7,906,864	93%

Illinois 2013				Kentucky 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992	5,050,157	1,469,935	29%	1992			
1994	3,106,566	1,346,813	43%	1994			
1996	4,311,391	2,113,839	49%	1996			
1998	3,394,521	1,952,599	58%	1998			
2000	4,742,123	3,001,753	63%	2000			
2002	3,538,883	2,556,665	72%	2002			
2004	5,274,322	3,905,283	74%	2004			
2006	3,486,671	2,968,217	85%	2006			
2008	5,523,051	4,802,142	87%	2008	1,826,508	1,754,525	96%
2010	3,729,989	3,530,491	95%	2010	1,356,468	1,371,391	101%
2012	5,242,014	5,175,513	99%	2012	1,797,212	1,798,505	100%

Michigan 2013				Montana 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992				1992	410,611	35,659	9%
1994				1994	350,387	35,594	10%
1996	3,848,844	3,607,184	94%	1996	407,083	77,832	19%
1998	3,036,886	3,147,777	104%	1998	331,551	122,047	37%
2000	4,232,501	4,383,047	104%	2000	410,986	276,739	67%
2002	3,177,565	3,301,336	104%	2002	331,321	245,049	74%
2004	4,839,252	5,031,387	104%	2004	450,445	315,033	70%
2006	3,801,256	3,901,267	103%	2006	406,505	382,997	94%
2008	5,001,766	4,856,418	97%	2008	491,960	480,567	98%
2010	3,226,088	3,182,562	99%	2010	360,341	371,831	103%
2012	4,730,961	4,582,529	97%	2012	484,048	504,880	104%

Table A5: Voter File Descriptive Statistics (Cont.)

Missouri 2005				Missouri 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992	2,391,565	398,191	17%	1992	2,391,565	232,112	10%
1994	1,775,116	534,331	30%	1994	1,775,116	260,656	15%
1996	2,158,065	964,818	45%	1996	2,158,065	489,638	23%
1998	1,576,857	996,740	63%	1998	1,576,857	426,518	27%
2000	2,359,892	2,110,893	89%	2000	2,359,892	1,323,351	56%
2002	1,877,620	1,776,931	95%	2002	1,877,620	1,182,202	63%
2004	2,731,364	1,882,308	69%	2004	2,731,364	2,030,647	74%
2006				2006	2,128,459	1,869,651	88%
2008				2008	2,925,205	2,603,503	89%
2010				2010	1,943,899	1,818,660	94%
2012				2012	2,757,323	2,740,083	99%
New Jersey 2013				Nevada 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992				1992			
1994				1994			
1996				1996	464,279	11,550	2%
1998				1998	435,790	211,405	49%
2000				2000	608,970	352,727	58%
2002				2002	504,079	349,178	69%
2004	3,611,691	3,666,375	102%	2004	829,587	596,511	72%
2006	2,250,070	2,122,133	94%	2006	582,572	478,978	82%
2008	3,868,237	3,678,654	95%	2008	967,848	838,029	87%
2010	2,121,584	2,044,172	96%	2010	721,404	681,060	94%
2012	3,640,292	3,019,452	83%	2012	1,014,918	1,005,143	99%
New York 2013				Oklahoma 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992				1992			
1994				1994			
1996				1996			
1998				1998			
2000				2000	1,234,229	522,770	42%
2002	4,579,078	4,240,053	93%	2002	1,035,620	582,652	56%
2004	7,391,249	7,146,506	97%	2004	1,463,758	982,306	67%
2006	4,490,053	4,716,949	105%	2006	926,462	773,520	83%
2008	7,640,640	7,048,787	92%	2008	1,462,661	1,303,260	89%
2010	4,658,586	4,686,242	101%	2010	1,034,767	978,782	95%
2012	7,074,723	6,958,083	98%	2012	1,334,872	1,326,492	99%
Oregon 2013				Pennsylvania 2013			
Election	Official Turnout	On Voter File	Percentage on File	Election	Official Turnout	On Voter File	Percentage on File
1992				1992			
1994				1994			
1996				1996			
1998				1998			
2000				2000			
2002				2002			
2004				2004	5,769,590	5,469,295	95%
2006	1,379,475	1,117,256	81%	2006	4,096,077	3,742,646	91%
2008	1,827,864	1,564,164	86%	2008	6,012,692	5,142,469	86%
2010	1,453,548	1,355,670	93%	2010	3,987,551	3,404,124	85%
2012	1,789,270	1,766,212	99%	2012	5,742,040	2,932,947	51%
Rhode Island 2013							
Election	Official Turnout	On Voter File	Percentage on File				
1992							
1994							
1996							
1998							
2000							
2002							
2004							
2006							
2008	471,766	437,012	93%				
2010	342,290	333,412	97%				
2012	446,049	445,966	100%				

2 Robustness of RD Estimation

In this section, we describe the robustness of our results to alternative specifications along three dimensions: functional form, inclusion or exclusion of controls, and the width of the window around the eligibility cutoff. The results presented in Tables 4 and 5 of the main analysis are generated by a model that includes first-order polynomial (linear) trends, controls for lagged voted totals, and a 365-day window on either side of the cutoff. The reasons this specification was chosen over alternatives were three-fold. First, controlling for lagged vote totals accounts for two main sources of heterogeneity in vote totals: day-of-the-week and season effects. Second, a 365-day window was chosen for the practical reason that one-year lagged vote totals are all "treated," whereas the lag for the 366th day after the cutoff is "untreated." Finally, we chose a first-order polynomial because after controlling for lagged vote totals, there remained no apparent curvilinear relationship with time.

In the tables below, we explore the sensitivity of our results to these specification choices. Table A6 considers the CACE of voting in 2008 on voting in 2012 and Table A7 does the same for the effect of voting in 2006 on voting in 2010. The rows of the tables correspond to functional forms: the difference-in-means model simply compares average vote totals on either side of the cutoff, whereas the polynomial models employ time trends of increasing flexibility to estimate the change in behavior at the cutoff. The columns of the table show windows of increasing size: from 3 months on either side of the cutoff to two years. Finally, the top and bottom panels present the results excluding or including controls for lagged vote totals, respectively.

Table A6: Robustness of Meta-Analytic Estimates (2008 on 2012)

	No additional controls							
	90 Days	180 Days	270 Days	365 Days	455 Days	545 Days	635 Days	730 Days
Difference-in-Means	0.132	0.132	0.125	0.110	0.124	0.131	0.126	0.120
	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
First-order Polynomial	0.068	0.122	0.137	0.148	0.115	0.111	0.125	0.136
	(0.012)	(0.008)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Second-order Polynomial	-0.028	0.057	0.098	0.123	0.155	0.144	0.118	0.108
	(0.019)	(0.013)	(0.010)	(0.008)	(0.007)	(0.007)	(0.007)	(0.006)
Third-order Polynomial	-0.063	-0.016	0.044	0.063	0.073	0.126	0.129	0.128
	(0.032)	(0.021)	(0.015)	(0.013)	(0.012)	(0.009)	(0.008)	(0.007)
	Controls for lagged vote totals							
	90 Days	180 Days	270 Days	365 Days	455 Days	545 Days	635 Days	730 Days
Difference-in-Means	0.108	0.118	0.124	0.119	0.118	0.117	0.110	0.103
	(0.005)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
First-order Polynomial	0.058	0.089	0.101	$\boxed{0.117}$	0.118	0.121	0.130	0.136
	(0.010)	(0.007)	(0.005)	$\boxed{(0.005)}$	(0.004)	(0.004)	(0.004)	(0.003)
Second-order Polynomial	0.016	0.055	0.074	0.083	0.102	0.104	0.099	0.103
	(0.015)	(0.010)	(0.008)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)
Third-order Polynomial	0.012	0.035	0.050	0.056	0.060	0.086	0.095	0.097
	(0.022)	(0.015)	(0.012)	(0.010)	(0.010)	(0.009)	(0.009)	(0.007)

Entries are estimated CACEs, with standard errors in parentheses.
Columns refer to the bandwidth around the cutoff.
Boxed estimate is used in the main analysis.

Table A7: Robustness of Meta-Analytic Estimates (2006 on 2010)

	No additional controls							
	90 Days	180 Days	270 Days	365 Days	455 Days	545 Days	635 Days	730 Days
Difference-in-Means	0.121	0.116	0.095	0.071	0.087	0.092	0.087	0.076
	(0.010)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
First-order Polynomial	0.073	0.111	0.140	0.147	0.103	0.092	0.103	0.117
	(0.024)	(0.014)	(0.011)	(0.010)	(0.009)	(0.008)	(0.008)	(0.007)
Second-order Polynomial	-0.025	0.043	0.060	0.112	0.160	0.159	0.119	0.098
	(0.047)	(0.028)	(0.020)	(0.015)	(0.013)	(0.012)	(0.012)	(0.011)
Third-order Polynomial	-0.019	0.027	0.081	0.004	0.052	0.101	0.130	0.157
	(0.071)	(0.040)	(0.034)	(0.029)	(0.023)	(0.018)	(0.018)	(0.016)
	Controls for lagged vote totals							
	90 Days	180 Days	270 Days	365 Days	455 Days	545 Days	635 Days	730 Days
Difference-in-Means	0.097	0.092	0.079	0.060	0.062	0.059	0.050	0.038
	(0.010)	(0.007)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
First-order Polynomial	0.061	0.090	0.112	$\boxed{0.119}$	0.098	0.091	0.098	0.106
	(0.023)	(0.014)	(0.011)	$\boxed{(0.009)}$	(0.008)	(0.008)	(0.007)	(0.007)
Second-order Polynomial	-0.006	0.038	0.046	0.086	0.125	0.129	0.099	0.089
	(0.044)	(0.026)	(0.019)	(0.015)	(0.013)	(0.012)	(0.011)	(0.011)
Third-order Polynomial	0.002	0.032	0.077	0.010	0.042	0.067	0.123	0.132
	(0.065)	(0.038)	(0.032)	(0.026)	(0.022)	(0.018)	(0.018)	(0.017)

Entries are estimated CACEs, with standard errors in parentheses.
Columns refer to the bandwidth around the cutoff.
Boxed estimate is used in the main analysis.

As Tables A6 and A7 show, our results are quite robust to specification. With the exception of the third-order polynomial (whose estimates are highly sensitive to points at the edges of the window) most estimates of the effect of 2008 on 2012 fall in the 0.10 to 0.12 range. The estimates of 2006 on 2010 are somewhat more variable, with most falling between 0.07 to 0.15. The estimates tend to be more precise the larger the window, the less flexible the functional form, and when controls are included.

3 Evidence that Primary and General Elections have Different Downstream Consequences

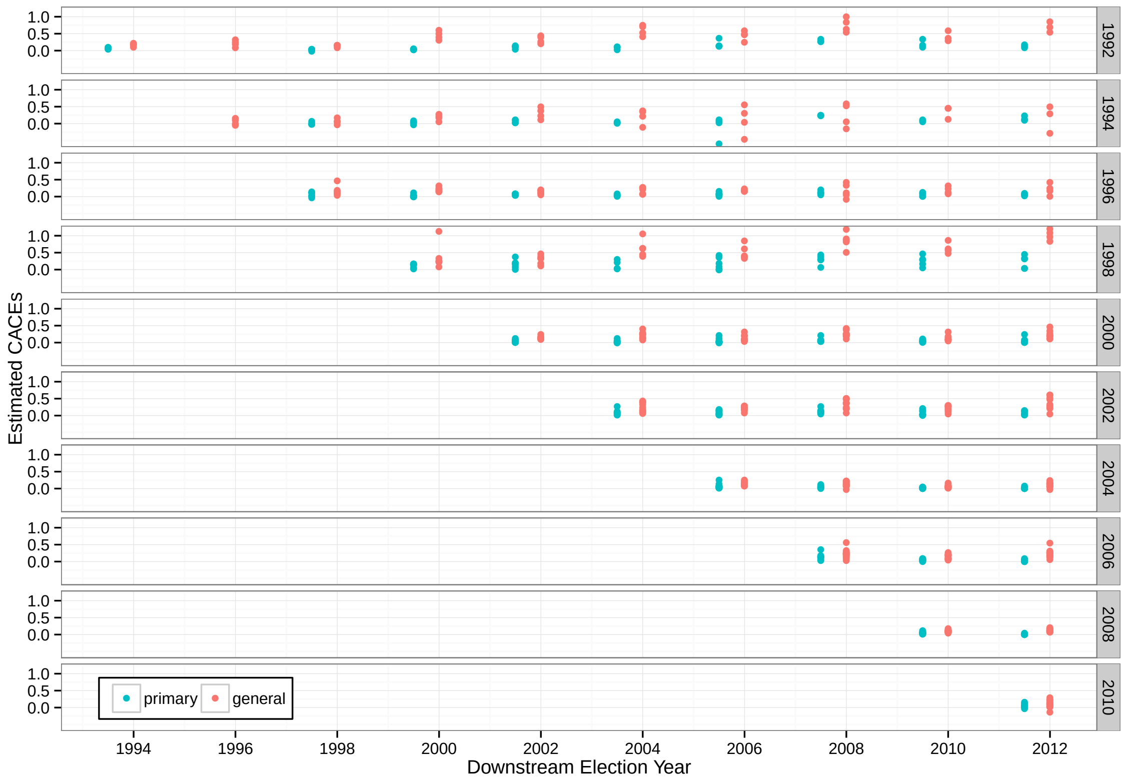
One hypothesis suggested by the downstream analysis of the three GOTV experiments is that the habits generated by voting depend on the type of upstream election in which they are formed.

The estimated CACEs appear to be larger in magnitude when the upstream and downstream elections are of the same type. For example, following encouragement to vote in the August 2006 primary election, the effect among compliers was much larger in subsequent August elections (0.135, 0.126, 0.089) than in subsequent November elections (0.108, 0.009, 0.043, 0.011). In principle, the discontinuity analysis offers an opportunity to confirm this "like elections" hypothesis.

This hypothesis is partially confirmed by an analysis of the downstream effects of eligibility to vote in upstream general elections on downstream primary and general elections. Habits formed by eligibility to vote in upstream general elections have larger impacts on downstream general elections than on downstream primary elections. Across the 384 general-on-general election pairs (the pairs reported in Table 6 of the main text), the average estimated CACE is 0.101 with a standard error of 0.007. Among the 335 general-on-primary pairs, the average estimate is 0.020, with a standard error of 0.004.

A visual representation of the "sawtooth" pattern of estimated CACEs is presented in Figure A2. Each point represents the estimated CACE for an upstream-downstream election pair in a single state. The series for each upstream year is plotted in a separate row – the tendency for primaries to be associated with weaker habit effects is apparent across all 10 upstream years. This finding lends credence to the hypothesis that habits formed in the course of voting are more strongly felt in downstream elections that are of the same type as the upstream election.

Figure A2: Downstream Effects of General Election Eligibility, by Upstream Election Year



We conducted the same analysis as above, using primary elections as the upstream election. Doing so introduced a new complication – while general election dates are the same nationwide, primary election dates vary state-to-state.[3] Some states hold primaries in the Spring while others wait until early Fall. If, as in the general election analysis, we were to use a 365-day window on either side of the eligibility cutoff, we would run into the difficultly that the general election discontinuity would be fast approaching for some states – we might misattribute some of the general election effect to the primary election. As a result, we restrict ourselves to a 60-day window, which is free from this contamination. Our estimates are accordingly less precise.

However, inducement to vote in primary elections appears to have negligible effects on downstream voting in either primary or general elections. The average estimated primary-on-general CACE is 0.0086 (SE = 0.0236), while the average estimated primary-on-primary CACE is 0.0068

[3]Thankfully, in our collection of states, primaries were always held on the same date for both major parties.

(SE=0.0010). These results differ from the evidence from the experiments, according to which voting in primary elections has large effects on subsequent behavior. We surmise that the individual characteristics of the compliers may explain this divergence. First, very few eligible 18-year-olds vote in primary elections, meaning that compliers account for a small sliver of the eligible population. Second, 18-year-olds who do comply, i.e., vote in the upstream primary when eligible, are likely to be very high-propensity voters. Voting in primary elections may not be habit-forming for these compliers precisely they have already acquired the habit somehow else.

Taken together, the suite of evidence we have assembled provides suggestive support for the hypothesis that habits generated in one context are more strongly felt in more similar future contexts. However, caution is still warranted. Each upstream encouragement to vote, whether experimental or naturally-occurring, generates its own distinctive cohort of compliers. The differences in electoral context could be driving the "like election" results, but differences between cohorts of compliers could as well.

4 Estimating the Effect of Eligibility on Campaign Contact: Evidence from a Special Tabulation of the ANES

We argued in the main text that the estimated habit-forming effects of voting are too large to be explained solely by increased campaign contact. In this section, we bolster that claim with survey evidence. We obtained a restricted version of the American National Election Survey Cumulative File that included the respondents' birthdates, so we are able to employ largely the same empirical strategy as in the main regression discontinuity analyses.[4] These data enable us to test whether those who in a previous election were just-eligible to vote were more likely to subsequently report mobilization activity than their counterparts, who were just-ineligible to vote. The ANES question asks "The political parties try to talk to as many people as they can to get them to vote for their candidate(s). Did anyone from one of the political parties call you up or come around and talk

[4]This confidential version of the file was approved by the ANES Board of Overseers and the Columbia University IRB.

to you about the campaign?" Following the same approach as the RD analysis presented in the main text, the model is specified using the following reduced form[5] equation, which is identical to equation A7, except that the unit of analysis is the individual survey respondent:

$$\text{Campaign Contact} = \beta_0 + \beta_1 Z + \beta_2 T + \beta_3 Z * T + \epsilon \tag{A12}$$

The parameter of interest, β_1, is the effect of eligibility to vote at the point of discontinuity. The environmental explanation implies a positive effect of eligibility on subsequent mobilization. We present results according to both 365-day and 180-day windows.

Table A8: Effect of Voting-Age Eligibility on Downstream Campaign Contact

	365-day Window		180-day Window	
1996 Eligibility	−0.136	(0.084)	−0.120	(0.124)
1998 Eligibility	−0.081	(0.086)	−0.007	(0.122)
2000 Eligibility	−0.005	(0.117)	0.210	(0.176)
2002 Eligibility	−0.149	(0.115)	−0.229	(0.184)
2004 Eligibility	0.203	(0.122)	0.258	(0.198)
2006 Eligibility	0.047	(0.114)	0.275	(0.177)
2008 Eligibility	−0.111	(0.149)	−0.042	(0.243)
2010 Eligibility	−0.123	(0.159)	−0.410	(0.253)
Meta-analysis Estimate	−0.053	(0.039)	−0.000	(0.059)
95% Confidence Interval	[−0.129, 0.024]		[−0.116, 0.116]	

Table A8 shows the effects of eligibility for each general election between 1996 and 2010 on campaign contact. The final row of the table pools the estimates across all elections using fixed-effects meta-analysis. According to the 365-day window, the estimated average effect of eligibility on self-reported mobilization activity is weakly negative, though not statistically significant. The just-eligibles are 5.3 percentage points *less* likely to be exposed to mobilization activity. According to the 180-day window, the average effect is exactly zero, with a confidence interval extending from

[5]Because the ANES data do note contain a reliable measure of whether treated subjects voted in the election for which they were "just-eligible," we fall back on the intent-to-treat (ITT) effect of eligibility on downstream consequences, also referred to as the reduced-form effect. Since the CACE is just a rescaled version of the ITT, this modeling approach does not impair our ability to assess whether the quasi-experimental encouragement to vote affects subsequent exposure to mobilization activity.

negative 11.6 percentage points to percentage 11.6 percentage points. Taken together, the evidence suggests that mobilization by political campaigns is unlikely to account for the apparent persistence in voting patterns.

5 Follow-up to the 2007 GOTV Experiment

As described in the "Downstream Results from Three GOTV Field Experiments" section of the main text, a follow-up to the 2007 social pressure experiment was conducted among the 27,138 subjects in either "Self" condition of the original experiment. Of these, a random 5,900 subjects were sent an additional "Self" mailer just prior to the November 2008 election that indicated whether the subject had indeed voted in the 2007 election. The refresher mailer had a no impact on voting behavior in November 2008: the estimated treatment effect was 0.002 with a cluster-robust standard error of 0.005.

However, we can take advantage of this follow-up experiment to assess the plausibility of the exclusion restrictions we invoked when studying the downstream effects of the original, quite effective, intervention. The exclusion restriction would be threatened if subjects continued to vote at high rates because they recall the 2007 social pressure mailing rather than because of the act of voting itself. We can indirectly assess whether this is the case by stimulating the recall of the original mailer with a followup mailer. Table A9 shows the downstream effects of voting in 2007 on voting in future elections, broken up by treatment group in the second round experiment. The estimated coefficients are never significantly different from one another, suggesting that the follow-up mailer does not rekindle memories of the original intervention.

Table A9: CACE of 2007 Upstream Vote on Downstream Voting Behavior, by Follow-up Condition

	Nov 2008 General	Aug 2010 Primary	Nov 2010 General	Feb 2012 P. Primary	Aug 2012 Primary	Nov 2012 General
Shown Vote + Recontact	0.119	0.041	0.167	-0.060	-0.204	0.017
	(0.091)	(0.130)	(0.127)	(0.116)	(0.136)	(0.120)
Shown Vote	0.088	0.090	0.115	-0.142	-0.108	0.130
	(0.052)	(0.076)	(0.075)	(0.071)	(0.078)	(0.068)

Robust standard errors clustered at household level in parentheses.

References

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–455.

Aronow, Peter M. 2013. "Model Assisted Estimation of Average Causal Effects under Generalized Noncompliance." *Unpublished Manuscript* .

Aronow, Peter M. and Donald P. Green. 2013. "Sharp Bounds for Complier Average Potential Outcomes in Experiments with Noncompliance and Incomplete Reporting." *Statistics and Probability Letters* 83(3):677–679.

Conley, Timothy G., Christian B. Hansen and Peter E. Rossi. 2012. "Plausibly Exogenous." *Review of Economics and Statistics* 94(1):260–272.

Meredith, Marc. 2009. "Persistence in Political Participation." *Quarterly Journal of Political Science* 4(3):187–209.

Rubin, Donald B. 1990. "Formal Modes of Statistical Inference for Causal Effects." *Journal of Statistical Planning and Inference* 25(25):279–292.