

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

Alexander Coppock Columbia University
Donald P. Green Columbia University

Field experiments and regression discontinuity designs test whether voting is habit forming by examining whether a random shock to turnout in one election affects participation in subsequent elections. We contribute to this literature by offering a vast amount of new statistical evidence on the long-term consequences of random and quasi-random inducements to vote. The behavior of millions of voters confirms the persistence of voter turnout and calls attention to theoretically meaningful nuances in the development and expression of voting habits. We suggest that individuals become habituated to voting in particular types of elections. The degree of persistence appears to vary by electoral context and by the attributes of those who comply with an initial inducement to vote.

Among the most robust empirical generalizations in political science is the observation that individual differences in voter turnout rates tend to persist over time. Scholars who have tracked voter turnout in successive elections have for decades observed that voting in one election strongly predicts voting in subsequent elections (Brody and Sniderman 1977) and that generations display persistent voter turnout patterns they acquire early in adulthood (Franklin 2004; Plutzer 2002).

What accounts for persistent interpersonal differences in turnout? One answer is that intrinsic motivation to vote endures over time. Long-standing attitudes and orientations, such as party identification or interest in politics, express themselves election after election (Milbrath 1965; Verba and Nie 1972). Social influences may also be stable across a series of elections: Certain voters may be continually mobilized by campaigns and members of their social network (Huckfeldt and Sprague 1992; Rosenstone and Hansen 1993). An alternative hypothesis offered by Green and Shachar (2000) and Gerber, Green, and Shachar (2003) is that the act of voting is itself habit forming. People who vote become accustomed to voting

and perhaps even acquire a taste for it. The act of voting itself increases the probability of voting in future elections.

Although complementary, these hypotheses may have quite different empirical implications. If voting is not only a recurrent manifestation of enduring psychological or situational factors but also reflects the causal influence of prior voting, variations in the current political environment may have long-lasting effects. For example, a drab, uncontested election may attract few voters, thereby disrupting voting habits and in turn lowering turnout in the next election cycle (Franklin and Hobolt 2011). If voting were solely a matter of intrinsic motives and external mobilization, a drab election need not depress turnout in subsequent elections; this effect would occur only if the current election campaign were to diminish attitudes conducive to voting or reduce the vigor with which mobilization activities occur in the future. Although finding that exogenous variations in past voter turnout affect subsequent turnout is not sufficient to establish the role of habit formation, this empirical pattern nevertheless constitutes an important prediction of the habit hypothesis, and the first order of business for any

Alexander Coppock is a Doctoral Candidate, Department of Political Science, Columbia University, 420 W. 118th Street, New York, NY 10027 (a.coppock@columbia.edu). Donald P. Green is Professor of Political Science, Columbia University, 420 W. 118th Street, New York, NY 10027 (dpg2110@columbia.edu).

The authors are grateful to Columbia University, which funded components of this research but bears no responsibility for the content of this report. This research was reviewed and approved by the Institutional Review Board of Columbia University (IRB-AAAM9102). We are grateful to Bradley Spahn for providing voter mobility data, to the secretaries of state and boards of elections from the 17 states whose voter history data are analyzed herewithin for providing a valuable public service, and to Peter Aronow, Lindsay Dolan, Winston Lin, participants in the Study of Development Strategies Seminar at Columbia University, and three anonymous reviewers for helpful comments. Replication code and data for all analyses are available at <http://dx.doi.org/10.7910/DVN/ALZVAW>.

American Journal of Political Science, Vol. 60, No. 4, October 2016, Pp. 1044–1062

©2015, Midwest Political Science Association

DOI: 10.1111/ajps.12210

investigation of habit is to establish whether random or as-if random shocks to voter turnout persist over time.

Prior observational, experimental, and quasi-experimental research has, on the whole, supported the hypothesis that exogenous shocks to voting persist over time. Franklin and Hobolt (2011) demonstrate that survey respondents in 27 European countries vote at significantly different rates depending on whether the first election in which they were eligible to vote was a European Parliament election or a national election. Consistent with the habit hypothesis, Franklin and Hobolt find that turnout is significantly lower among those who first became eligible to vote in the run-up to a European Parliament election, which tends to draw low turnout, as opposed to a national election. Atkinson and Fowler (2014) find that local religious festivals in Mexico depress electoral turnout and that, consistent with the habit hypothesis, turnout is depressed in the following election as well. Denny and Doyle (2009) analyze a long-term British panel study, using the number of times respondents moved in young adulthood as an instrumental variable. They find that voting in prior national elections raises subsequent turnout significantly, even after controlling for individual fixed effects and an extensive set of background attributes. Evidence of persistent voting shocks comes also from the instrumental variables analysis presented by Green and Shachar (2000), who use the perceived closeness of the election and ideological distance between the candidates to identify the effects of past turnout on subsequent turnout in the 1972–76 and 1992–96 American National Election Studies (ANES) panel studies.

Unlike these observational studies, which attempt to isolate an as-if random inducement to vote at one point in time and trace whether turnout is elevated in subsequent elections, experimental studies begin with a truly random inducement. The first large-scale experimental study to trace the so-called “downstream effects” (Green and Gerber 2002) of a get-out-the-vote (GOTV) drive examined encouragements to vote by canvassers or direct mail in the 1998 midterm elections and found that they elevated turnout both in that election and in the municipal elections held one year later. Gerber, Green, and Shachar (2003) report that approximately half of the turnout effect observed in 1998 persisted through 1999. Michelson (2003), on the other hand, finds no evidence that a successful GOTV campaign conducted prior to a 2001 municipal election in a California farming community had persistent effects a year later in a midterm election. That experiment, however, was relatively small, and its power to detect persistent effects was limited. Much better powered are the 15 experiments presented by Bedolla and Michelson (2012, 178), which track the

enduring effects of GOTV campaigns directed at minority voters in California over a series of primary and general elections from 2006 to 2008. Their overall assessment is that voting in an “upstream” election increases the probability of voting in a subsequent “downstream” election by 23 percentage points. Further support for the habit hypothesis comes from a GOTV experiment conducted in the context of a UK general election in 2005; tracking the downstream effects of canvassing and phone calls in a 2006 municipal election, Cutts, Fieldhouse, and John (2009) find that approximately half of the initial mobilization effect persisted a year later. However, Hill and Kousser (2015) describe a direct mail GOTV campaign that had small but significant turnout effects in a primary election yet did not increase general election participation.

One crucial assumption underlying these downstream analyses is that the GOTV campaign that raised turnout in the initial election does not itself increase turnout a year or more later (except via its indirect influence on turnout in the initial election). Ordinarily, that assumption is plausible—as Gerber, Green, and Shachar (2003) point out, one would scarcely expect voter mobilization activities to be effective if launched a year prior to an election. A possible exception, however, is GOTV campaigns that employ what Gerber, Green, and Larimer (2008) call “social pressure”: messages that forcefully assert the norm of doing one’s civic duty and promise that compliance with this norm will be verified using public records and publicized. A series of experiments (Abrajano and Panagopoulos 2011; Davenport 2010; Gerber, Green, and Larimer 2010; Panagopoulos 2010; Mann 2010; but see Matland and Murray 2012) has shown these messages to be much more influential than conventional GOTV appeals, perhaps because they are especially memorable. Evidence that social pressure effects endure, therefore, is not necessarily evidence of habit formation. Nevertheless, it is worth noting that social pressure messages do seem to have significant downstream consequences for turnout in subsequent elections.¹ Davenport et al. (2010) trace the effects of six experiments over time and find in some cases that treatment effects persist eight years after the initial intervention.

An alternative to the observational and experimental designs described above is regression discontinuity analysis, which capitalizes on the fact that eligibility requirements set in motion different voting trajectories among people who were born only a few days apart. The use of

¹The same persistence also holds for the effects of mailings from the secretary of state saying that ballots are secret (Gerber et al. 2014).

regression discontinuity to study voting habits was pioneered by Meredith (2009), who compared those who were just over or under 18 at the time of the 2000 election. He found that Californians who were eligible to vote in 2000 were significantly more likely to vote in the 2004 presidential election. The size of this four-year habit effect, however, is on the order of 7 percentage points, which is smaller than would be expected from the body of research based on downstream analyses of GOTV field experiments. The advantage of Meredith's method is that it may be applied to any state that provides voter turnout records and detailed birthdate information in its public voter file. Dinas (2012) applies a similar regression discontinuity design to a panel survey stretching from 1965 to 1997, finding large and persistent effects of voting eligibility on downstream voting behavior. Although the investigation of habit formation through regression discontinuity analysis is limited to young voters for whom the "assignment" of eligibility is not strictly random, this approach has the advantage of sidestepping certain statistical problems that arise when experiments are tracked over time.²

Our contribution to the literature on voting habit is threefold. First, we offer an array of new evidence based on downstream experiments and regression discontinuities. The sheer volume of fresh statistical evidence brings new precision to the study of habit, qualifying and in some cases overturning previous interpretations. The overall pattern leaves little doubt that voting is habit forming. Second, we make the case that voting habits are more nuanced than previous accounts have recognized. Our results suggest, for example, that the strength and persistence of voting habits vary by electoral context and by the attributes of those who comply with an initial inducement to vote. In particular, habits are more strongly expressed among those induced to vote in low-salience general elections and among those who are more residentially stable. Finally, we show that voting habits formed early in life can persist for decades: The causal chain of events set in motion by an eligibility discontinuity in 1992 has detectable effects on voting behavior in 2012. The sections that follow describe habit formation formally, lay out the conditions under which its effects may be identified, present results from experiments and discontinuities, and discuss the implications for future research.

²As explained below, these issues are weak instruments (i.e., the randomized intervention does not strongly affect turnout in the initial election) and file-drawer problems (i.e., only a subset of experiments, those that show strong positive effects, are tracked over time, which implies that disturbances affecting outcomes tend on average to favor the treatment group).

Detecting Habit Formation Using Experiments and Discontinuities

In the context of electoral participation, the concept of habit implies that if two people whose psychological propensities to vote were identical should happen to make different choices about whether to go to the polls on Election Day, these behaviors will alter their chances of voting in the next election. In other words, holding preexisting individual and environmental attributes constant, merely going to the polls increases one's chance of returning. The claim is not simply that individual differences in voting propensity persist over time, which is apparent from simple cross-tabulations of voting behavior among respondents in panel studies. Rather, the hypothesis is that when one votes, the propensity to vote in future elections increases. To formalize this hypothesis and show how an experimental stimulation of turnout can be used to isolate habit formation, we must construct a model that explicitly allows for unobserved heterogeneity among individuals.

As Gerber and Green (2012, chap. 6) point out, the study of voting habits is formally analogous to the study of experiments with two-sided noncompliance. The estimand of interest is the average effect of voting in an upstream election (V_1) on voting in a downstream election (V_2). We cannot randomly assign voting in the upstream election; at most, we can randomly assign some sort of encouragement, which may take the form of a GOTV appeal or legal eligibility to vote. Call this encouragement $Z \in \{0, 1\}$, where 1 indicates that encouragement is offered and 0 otherwise. The identification question is, What can we learn from a design in which Z is randomly assigned and Z , V_1 , and V_2 are observed? What assumptions must be invoked along the way?

Crucially, we assume that Z is independent of potential outcomes. For the experiments, independence is ensured by design; for the discontinuities, we assume that assignment is as-if randomly assigned at the point of discontinuity. In addition to independence, we must assume excludability, which requires that Z have no impact on V_2 except through its influence on V_1 . Finally, we must invoke the noninterference assumption, which states that units' potential outcomes do not respond to others' treatment assignments.

We define the Average Upstream Treatment Effect (AUTE) as $E[V_{1i}(1)] - E[V_{1i}(0)]$, where $V_{1i}(1)$ and $V_{1i}(0)$ refer to unit i 's treated and untreated potential outcomes, respectively. The Average Downstream Treatment Effect (ADTE) is analogously defined as $E[V_{2i}(1)] - E[V_{2i}(0)]$. Our estimand of interest is the complier average causal effect (CACE), defined as $E[V_{2i}|V_{1i} =$

1] - $E[V_{2i}|V_{1i} = 0]$, the effect of voting in an upstream election on downstream participation, among a subset of subjects, the compliers. Compliers are those who vote in the upstream election if and only if they receive the encouragement. An estimator of the CACE is given in Equation (1).

$$\widehat{CACE} = \frac{\hat{E}[V_{2i}|Z_i = 1] - \hat{E}[V_{2i}|Z_i = 0]}{\hat{E}[V_{1i}|Z_i = 1] - \hat{E}[V_{1i}|Z_i = 0]} = \frac{\widehat{ADTE}}{\widehat{AUTE}} \quad (1)$$

We will employ two versions of this estimator to address the specific features of our two research designs. In the case of the experiments, we will estimate the CACE via two-stage least squares (Angrist, Imbens, and Rubin 1996). The discontinuities present a challenge because public voter files do not list those who are eligible to vote but have not registered. For this reason, we employ the total number of votes cast by a birthdate cohort as the unit of analysis and construct a two-stage least squares estimator out of the following reduced-form equations:

$$\begin{aligned} \text{Downstream Votes Cast} = & \beta_0 + \beta_1 Z_j + \beta_2 T_j \\ & + \beta_3 Z_j * T_j \\ & + \beta_4 \text{Lagged Downstream}_j \\ & + \epsilon_j \end{aligned} \quad (2)$$

$$\begin{aligned} \text{Upstream Votes Cast} = & \alpha_0 + \alpha_1 Z_j + \alpha_2 T_j \\ & + \alpha_3 Z_j * T_j \\ & + \alpha_4 \text{Lagged Downstream}_j + \eta_j \end{aligned} \quad (3)$$

where Z_j is an indicator for eligibility to vote in the upstream election, T_j is a running variable indicating the number of days between a birthdate and the eligibility cutoff, and j indexes birthdate cohorts. We include $\text{Lagged Downstream}_j$ (the total votes cast by birthdate cohort one year older) to account for seasonal and day-of-the-week birth trends. For full details of our identification strategy, including a discussion of the complications arising from the use of voter histories that are possibly measured with error, see the supporting information.

Downstream Results from Three GOTV Field Experiments

As mentioned earlier, the study of downstream effects requires careful case selection. If one selects from the pool of statistically significant AUTE estimates, one runs the risk of assembling a collection of lucky draws in which the balance of unobservables favors the treatment group over the control group. If these unobserved factors persistently favor the treatment group, one may obtain apparent habit

effects where none exist. In order to minimize this bias, the experiments in question should be large so that sampling variability plays a minor role, and they should have a strong track record of replicability so that there is no reason to think that they are drawn from the extremes of the sampling distribution. A further consideration is that the treatments should generate a large number of compliers; in other words, the intervention should have a sizable effect on voting in the election immediately following the treatment. Indeed, an important advantage of experiments over discontinuity analysis is that experiments can produce ample numbers of compliers in very low-salience elections, whereas eligibility discontinuity “encouragements” tend to be weak. Three experiments that satisfy these criteria are the social pressure studies reported by Gerber, Green, and Larimer (2008, 2010) and Sinclair, McConnell, and Green (2012), which are large, apply similar treatments, and produce substantial upstream turnout effects.

The 2006 Social Pressure Experiment

Setting. Gerber, Green, and Larimer (2008) report the results of an experiment conducted in Michigan prior to its August 2006 primary election, which featured few competitive contests and attracted only 18% of the registered electorate. The subject pool for their experiment comprised 180,002 households (people with the same last name and living at the same address) for which the voter file provided valid vote history in prior elections. These households were also restricted to people living in relatively populous blocks with few apartment buildings; the sample was further screened to exclude those with a high probability of voting absentee or in the Democratic primary. We study experimental participants who were present in the voter file as of 2013.

Treatments. Households were blocked according to their street address and assigned to one of five experimental groups. A control group received no mail, and the four treatment groups each received a different mailing. The “Civic Duty” treatment urged recipients to “Remember your rights and responsibilities as a citizen. Remember to vote.” The second mailing added to this civic duty baseline a mild form of social pressure: in this case, observation by researchers. Households receiving the “Hawthorne Effect” mailing were informed that their voting behavior would be examined by means of public records. The “Self” mailing exerted more social pressure by informing recipients that who votes is public information and listing the recent voting record of each registered voter in the household. The mailing informed voters that

after the primary election, “we intend to mail an updated chart,” filling in whether the recipient voted in the August 2006 primary. The fourth mailing, “Neighbors,” listed not only the household’s voting records but also the voting records of those living nearby. Like the Self mailing, the “Neighbors” mailing informed the recipient that “we intend to mail an updated chart” after the primary, showing whether members of the household voted in the primary and who among their neighbors had voted in the primary. The implication was that members of the household would know their neighbors’ voting records, and neighbors would know theirs. By threatening to “publicize who does and does not vote,” this treatment applied additional social pressure.

The 2007 Social Pressure Experiment

Setting. Gerber, Green, and Larimer (2010) conducted a similar experiment with a different subject pool in Michigan prior to the November 2007 municipal elections. In November 2007, 224 cities in the state of Michigan held elections, and the salience of these elections varied markedly. Some elections featured a range of offices and ballot measures, whereas others featured only minor and often uncontested offices.

Study Population. Subjects for this experiment were again voters from Michigan’s voter file. For purposes of random assignment, voters were grouped into households prior to assignment. Households with three or more registered voters were eliminated, as were households that received mailings in the 2006 experiment or resided in precincts with fewer than 100 registered voters. Unlike the 2006 study, subjects were not restricted based on their likelihood of voting absentee or on their partisan profile.

Treatments. Each household was randomly assigned to one of four groups. A control group received no mail, and the three treatment groups received a mailing that urged them to vote. The first treatment was the Civic Duty mailing used in the prior study. The other two mailings were variants of the Self mailing, with one mailing reporting past voting from 2005 and the other from 2006. Because both studies randomly assigned treatments to households rather than individuals, Tables 1 and 2 present robust standard errors clustered at the household.

Another noteworthy aspect of the Gerber, Green, and Larimer (2010) study is that a random sample of subjects assigned to the Self mailing received a follow-up mailing a year later (immediately prior to the 2008 general election) indicating whether they had voted in the 2007 election. This follow-up mailing was designed to

test whether rekindling memories of the initial mailing increased turnout in the 2008 general election and increased apparent downstream effects in that and subsequent elections. This aspect of the experiment, in effect, tests one possible violation of the exclusion restriction, namely, the enduring effects of remembering the initial treatment.

The 2009 Multilevel Experiment

Setting. Sinclair, McConnell, and Green (2012) conducted a study similar to the Gerber, Green, and Larimer (2010) experiment but used a more elaborate randomization scheme in order to measure intra-neighborhood and intra-household spillover effects. The setting was Illinois’ 5th Congressional District, which held a special federal election to fill a vacancy in April 2009.

Study Population. All subjects had registered to vote prior to 2006 and resided in households containing no more than three registered voters. Voters were included in the study only if their nine-digit zip code comprised 3–15 total households, at least two of which contained two voters. The fact that treatment probabilities vary by the number of voters in a household makes it necessary to control for household size.

Treatment. A single mailer was sent to one person in each household assigned to the treatment group. This mailing was patterned after the Self mailing from Gerber, Green, and Larimer (2010) except that it listed the subject’s participation in the previous two spring elections. The analysis below focuses solely on the effects of receiving the mailer directly as opposed to being in an untreated household; we ignore whether one’s neighbors receive mailings (which Sinclair, McConnell, and Green 2012 found to have no effect) and exclude indirectly treated individuals whose housemates received mailings. Standard errors take into account the fact that individuals are assigned to the control group as household clusters.

Upstream and Downstream Results

The first columns of Tables 1 and 2 report estimated average upstream treatment effects for each of the mailers on turnout in the election immediately following receipt of the mail. For the two types of mailings that were used in both experiments, Civic Duty and Self, the average upstream treatment effects are similar. In 2006, the Civic Duty mailing increased turnout by 1.8 percentage points, as compared to 1.4 percentage points in 2007. The Self mailing increased turnout by 5.0 percentage points in 2006 and by 4.5 and 5.0 percentage points in 2007,

TABLE 1 Immediate and Downstream Effects of 2006 Social Pressure Messages on Voter Turnout

Instrument	Instrumental Variable Estimates (Instrumented = Aug. 2006)									
	First Stage									
	Aug. 2006 Primary	Nov. 2006 General	Jan. 2008 P. Primary	Aug. 2008 Primary	Nov. 2008 General	Aug. 2010 Primary	Nov. 2010 General	Feb. 2012 P. Primary	Aug. 2012 Primary	Nov. 2012 General
Civic Duty n = 36, 903	0.018 (0.003)	0.041 (0.105)	0.278 (0.180)	-0.105 (0.186)	-0.078 (0.098)	0.067 (0.183)	-0.035 (0.151)	-0.023 (0.181)	-0.105 (0.187)	-0.173 (0.136)
Hawthorne n = 37, 005	0.025 (0.003)	-0.014 (0.077)	0.117 (0.130)	0.118 (0.127)	-0.109 (0.073)	0.166 (0.131)	-0.176 (0.115)	0.114 (0.130)	0.060 (0.131)	-0.001 (0.094)
Self n = 37, 011	0.050 (0.003)	0.050 (0.038)	0.133 (0.066)	0.169 (0.064)	-0.010 (0.035)	0.134 (0.066)	0.052 (0.055)	0.073 (0.066)	0.185 (0.065)	0.055 (0.047)
Neighbors n = 36, 893	0.083 (0.003)	0.128 (0.022)	0.146 (0.040)	0.125 (0.039)	0.017 (0.021)	0.122 (0.040)	0.047 (0.033)	0.072 (0.040)	0.061 (0.040)	-0.002 (0.029)
All Instruments		0.108 (0.021)	0.142 (0.036)	0.135 (0.036)	0.009 (0.019)	0.126 (0.037)	0.043 (0.030)	0.073 (0.036)	0.089 (0.036)	0.011 (0.026)
Control n = 184, 749	0.311 (0.001)									
Untreated Compliers		[0.845, 0.943]	[0.194, 0.362]	[0.274, 0.398]	[0.912, 1.034]	[0.419, 0.472]	[0.766, 0.986]	[0.330, 0.453]	[0.286, 0.469]	[0.806, 0.950]

Notes: Each cell represents a single regression. Robust standard errors clustered at the household level are in parentheses. The estimates in the "All Instruments" row are overidentified and are obtained using 2SLS. Numbers in brackets represent the minimum and maximum estimated untreated turnout rates among the four treatments' compliers.

TABLE 2 Immediate and Downstream Effects of 2007 Social Pressure Messages on Voter Turnout

Instrument	Instrumental Variable Estimates (Instrumented = Nov. 2007)								
	First Stage Nov. 2007 Municipal	Jan. 2008 P. Primary	Aug. 2008 Primary	Nov. 2008 General	Aug. 2010 Primary	Nov. 2010 General	Feb. 2012 P. Primary	Aug. 2012 Primary	Nov. 2012 General
Civic Duty n = 6, 815	0.014 (0.006)	0.180 (0.466)	0.642 (0.468)	-0.016 (0.326)	0.814 (0.522)	-0.101 (0.493)	0.248 (0.423)	-0.322 (0.539)	-0.042 (0.433)
Shown 2005 Vote n = 13, 592	0.050 (0.005)	0.373 (0.091)	0.218 (0.086)	0.066 (0.062)	0.121 (0.091)	0.077 (0.089)	-0.148 (0.086)	-0.036 (0.091)	0.140 (0.081)
Shown 2006 Vote n = 13, 546	0.045 (0.005)	0.298 (0.101)	0.120 (0.097)	0.128 (0.068)	0.032 (0.103)	0.181 (0.099)	-0.096 (0.094)	-0.235 (0.109)	0.066 (0.091)
All Instruments		0.336 (0.067)	0.183 (0.064)	0.092 (0.046)	0.095 (0.068)	0.119 (0.066)	-0.118 (0.063)	-0.128 (0.070)	0.104 (0.061)
Control n = 759, 964	0.282 (0.001)								
Untreated Compliers		[0.191, 0.525]	[-0.157, 0.247]	[0.838, 1.024]	[0.017, 0.425]	[0.591, 0.994]	[-0.001, 0.310]	[0.393, 0.779]	[0.755, 0.845]

Notes: Each cell represents a single regression. Robust standard errors clustered at the household level are in parentheses. The estimates in the "All Instruments" row are overidentified and are obtained using 2SLS. Numbers in brackets represent the minimum and maximum estimated untreated turnout rates among the three treatments' compliers.

depending on which past election's turnout was reported in the mailing. The Neighbors mailing, used only in 2006, generates even stronger effects, raising turnout by 8.3 percentage points. Each of these treatment effects is statistically significant at the .05 level or better. However, because the precision of the instrumental variables estimates grows with the size of the immediate treatment effect, much of what we glean from this downstream experiment derives from the Self and Neighbors treatments.

Tables 1 and 2 also report the instrumental variables regression estimates of the CACE using each of the mailings as instruments for voting in the election immediately following the intervention. Our interpretation focuses on the estimates reported in the All Instruments row, which represent an efficient summary of all of the treatments combined. Finally, the tables report the minimum and maximum estimates of the voter turnout rate among untreated compliers, which differ depending on which treatment category is used as an instrument.³ These turnout rates aid the interpretation of the estimated CACE because they indicate whether the apparent influence of habit is constrained by ceiling effects.

Looking first at Table 1, which reports results from the 2006 study, we see that the two-stage least squares (2SLS) estimate of the effect of voting in August 2006 on voting in November 2006 is 0.108, with a standard error of 0.021. This estimate is highly significant ($p < .001$) but also substantively larger than it may at first appear. The last row of Table 1 indicates that the lowest estimate of the voting rate among untreated compliers is 84%. This figure in conjunction with the estimated CACE implies that voting in August raised turnout among compliers from 84% to 95%, which is about as high as turnout can plausibly go. The next two elections were primary elections. In January 2008, Michigan held a presidential primary, which attracted relatively low turnout among compliers.⁴ Voting in the August 2006 primary raised turnout among compliers by 14.2 percentage points ($SE = 0.036$, $p < .001$). Similarly, the estimated CACE is a statistically significant 13.5 percentage points in the August 2008 primary, fully two years after the first election following the GOTV campaign. Tracking the estimated CACE across the full set of August elections shows a gradual pattern of decline, with an estimated CACE of 12.6 percentage points in 2010 and 8.9 percentage points in 2012, both of which remain statistically significant at the .05 level.

³For more detailed information on how this figure is calculated, see Aronow and Green (2013) and the supporting information.

⁴This election was controversial because Michigan moved its primary to an early date in defiance of the parties' wishes, and the parties reacted by threatening to reduce Michigan's representation in the party conventions.

By contrast, no significant effects are obtained for subsequent November elections. The overall pattern of results suggests that voting in August 2006 had a sizable effect on turnout in the general election a few months later but thereafter had strong and persistent effects only on other primary elections.

Table 2 reports the downstream effects of voting in the November 2007 municipal elections. The estimated CACE on voting in the presidential primary held two months later is sizable at 33.6 percentage points ($SE = 6.7$). The estimated CACE remains substantial for the August 2008 primary (18.3 percentage points, $SE = 6.4$) and exerts a marginally significant effect on turnout in the presidential election in November 2008 ($\widehat{CACE} = 0.092$, $SE = 0.046$), where turnout among compliers was 84% or higher, but after that, the estimated CACEs become equivocal.

Interestingly, the results in Table 2 are essentially unchanged when we restrict our attention to subjects who in 2007 received the Self mailing and were mailed a reminder with their updated vote history prior to the 2008 general election. The follow-up mailing had no direct effect whatsoever on voting in 2008, and the apparent downstream voting effects among the Self recipients who received the reminder are not consistently stronger than the corresponding downstream effects among those who did not receive the reminder.⁵ This finding bolsters the credibility of the exclusion restriction because it suggests that downstream voting is due to upstream behavior rather than the recollection of the mailers.

In sum, the short half-life of habit effects in the 2007 study contrasts with the persistent effects apparent from the 2006 study. In both cases, the GOTV campaign encouraged participation in a low-salience election. Encouragement to vote in an August primary had long-lasting effects on voting in subsequent primaries, especially subsequent August primaries. Municipal voting raised downstream turnout initially but thereafter had sporadic effects.

An instructive intermediate case comes from an experiment conducted prior to the 2009 special congressional election (Sinclair, McConnell, and Green 2012). Unlike the 2007 study, the focus here is on federal candidates; unlike the 2006 study, the context in 2009 is a general election featuring two-party competition; and unlike both prior studies, this one takes place in April.

⁵The direct treatment effect of the reminder was 0.2 percentage points, with a standard error of 0.5 percentage points. Downstream effects are estimated to be stronger among recontacted subjects in three of the six subsequent elections, and the difference in downstream effects is never significant. For full results of this follow-up experiment, please see the supporting information.

TABLE 3 Immediate and Downstream Effects of 2009 Social Pressure Messages on Voter Turnout

Instrument	First Stage Apr. 2009 Special	Instrumental Variable Estimates (Instrumented = Apr. 2009)					
		Feb. 2010 Primary	Nov. 2010 General	Feb. 2011 M. Primary	Apr. 2011 Municipal	Mar. 2012 Primary	Nov. 2012 General
Household Size = 1 n = 16,638 Untreated Compliers	0.049 (0.008)	0.478 (0.165)	0.399 (0.177)	0.306 (0.22)	0.462 (0.183)	0.446 (0.15)	0.236 (0.163)
Household Size = 2 n = 16,915 Untreated Compliers	0.035 (0.009)	-0.366 (0.322)	0.06 (0.263)	0.301 (0.328)	0.463 (0.276)	0.043 (0.257)	0.231 (0.233)
Household Size = 3 n = 5,086 Untreated Compliers	0.044 (0.018)	0.309 (0.422)	0.389 (0.437)	0.149 (0.487)	-0.223 (0.493)	-0.063 (0.437)	0.142 (0.409)
Pooled Estimate n = 38,639	0.043 (0.005)	0.303 (0.139)	0.303 (0.139)	0.285 (0.171)	0.403 (0.146)	0.311 (0.124)	0.226 (0.127)

Notes: Each cell represents a single regression. Robust standard errors clustered at the household level are in parentheses.

Thus, the question is whether voting in this election induces higher turnout in subsequent federal elections and spring elections. Table 3 indicates that compliers are substantially more likely to vote in both types of elections. Regarding spring and winter elections, the pooled results reveal significant increases in turnout among compliers in February 2010, April 2011, and March 2012. For example, compliers were 40.3 percentage points more likely to vote in April 2011 as a result of voting in April 2009 ($p < .01$). As for November elections, the effects are strong but gradually diminishing, with estimated CACEs of 0.303 in 2010 and 0.226 in 2012. The overall pattern suggests that the type of “upstream” election (general, primary, municipal, or special) plays an important role in habit formation. Habituation seems most apparent when downstream elections and the upstream election are of the same type. We now turn to evidence from regression discontinuities, which allow us to study the expression of habit across a wide spectrum of elections that vary in salience and institutional context.

Eligibility Discontinuities in 17 U.S. States

We obtained voter files from 17 states: Arkansas, Colorado, Connecticut, Iowa, Illinois, Florida, Kentucky, Michigan, Missouri, Montana, Nevada, New Jersey, New York, Oklahoma, Oregon, Pennsylvania, and Rhode Island. These files contain statewide voter eligibility and history data. The scope of our data collection beyond

these states was limited by state-specific disclosure policies. Some states (e.g., Ohio) do not provide birthdates; others (e.g., North Carolina) provide only age ranges on request; still others (e.g., California) provide vote history for no more than the most recent eight elections in which a voter has participated. We obtained the publicly available official records directly in all but one state. The exception was Michigan, whose file was provided by Practical Political Consulting.

One wrinkle associated with using official turnout records is that the “age” of the voter file—the number of years that have elapsed between the upstream election and the date when the voter file was assembled—may affect the results in substantively meaningful ways. A historical voter file is more likely to have lost track of then-young voters who moved out of the state (or moved without reregistering) in the intervening years. Because states no longer purge nonvoters from the rolls (Highton and Wolfinger 1998), restricting the subject pool does not necessarily imply that the estimated CACE is biased when estimated using an old voter file; however, the group of voters to whom the estimates pertain is weighted more heavily toward residentially stable compliers. A contemporaneous voter file contains a mix of residentially stable and residentially mobile people, and the CACE refers to this mixture of complier types. In order to assess the effect of voter file age on the estimated CACE, we leverage the fact that for two states, Florida and Missouri, we have voter files from both 2005 and 2013, which allows us to estimate the CACE for the same election pairs among voters with varying levels of residential mobility.

We employ a 365-day window on either side of the eligibility cutoff, which is the maximum window that allows for the inclusion of the Lagged Downstream covariate. As noted above, we employ a first-order polynomial (linear) functional form. For the sensitivity of our results to smaller windows and higher-order polynomials, see the supporting information.

A final concern with the discontinuity strategy is possibility of bias due to measurement error. The official voter lists maintained by secretaries of state might contain mistakes for a variety of reasons, including administrative errors, failure to purge former residents, and lost voter histories. For the purposes of the present investigation, it is voter mobility that presents the largest challenge. Our estimand is the CACE among the residents of a state at the time the voter file was generated. Consider two just-eligible voters, both of whom vote in the upstream election. The first moves out of state, whereas the second moves into the state. If the secretary of state does not remove the first voter from the list, our estimate of the average upstream treatment effect (AUTE) will be too large, causing us to underestimate the CACE. The second voter's history does not travel with her from out of state: We record her as not having voted in the upstream election. Our estimate of AUTE is therefore too small, and the resulting CACE estimate is too large.

There are two reasons to believe that the net bias is likely to be small. First, we obtained an estimate, by state, of the number of young people who voted in different states in the 2008 and 2012 elections.⁶ The number of movers is small in relation to the total size of the population near the cutoff. Further, the number of in-migrants to a state is generally similar to the number of out-migrants. Second, under the excludability assumption that the eligibility cutoff does not affect mobility, our estimates of the average downstream treatment effect (ADTE) are not biased. Since our hypothesis tests depend only on the ADTE and its associated sampling variability, all our inferences about the sign and significance of the CACE are unaffected by this type of measurement error. For further discussion of this issue, see the supporting information.

Results

We now analyze a series of eligibility discontinuities to assess whether voting shortly after turning 18 affects subsequent turnout. Tables 4 and 5 present downstream estimates and accompanying standard errors for each state.

⁶We thank Bradley Spahn for generously providing this tabulation.

In order to make the presentation manageable, we have grouped the relevant election years into four categories: (1) presidential on presidential, which assesses the effects of voting in one presidential election on the probability of voting four years later; (2) midterm on midterm, which tracks habit formation over a two-year period, starting with a midterm election; (3) presidential on midterm, which assesses habit formation over a two-year period, starting with a presidential election; and (4) midterm on presidential, which tracks habit formation over a two-year period, starting with a midterm election. This categorization allows the reader to assess the effects of compliance in two very different contexts, presidential and midterm elections, on turnout two and four years later.

Across all years and states, midterm elections attract a small percentage of registered just-eligible voters, typically less than 10%. Even presidential elections attract roughly one-third of the just-eligible electorate.⁷ Thus, by the standards of this age cohort, “midterm compliers” (those who vote if they are just-eligible to participate in a midterm election) are relatively high-propensity voters while “presidential compliers” are less so.

Table 4 shows how habits persist over two-year periods. Looking first at presidential-on-midterm effects, we see that all 54 estimates are positive. Using fixed-effects meta-analysis, we find the precision-weighted average estimate for the period 2008–10 to be 0.09, with a standard error of just 0.002. We see a gradual upward progression in these pooled estimates as we work backward in time, consistent with the hypothesis that more residentially stable people are less likely to have their habits disrupted. The effects of voter file age are evident from close inspection of Florida and Missouri prior to 2003, where the historical voter files tend to generate higher estimates of the CACE than the contemporaneous files. Regardless of whether we choose to focus on residentially stable voters, the magnitude of the habit effect is substantively large. A \widehat{CACE} of 0.09 in Missouri, for example, implies that casting a vote in 2008 increases the probability that a complier will cast a vote in the 2010 midterm election from 15.4% to 24.4%.

The pattern of midterm-on-presidential estimates argues strongly for habit formation as well. Here, 50 out of 54 estimates are positive. The average effect of turnout

⁷These figures are calculated directly from the voter files, so the turnout rate is conditional on registering to vote. Using voter file information, Meredith (2009) estimates the turnout rate of all 18- and 19-year-olds in the 2000 and 2004 elections at 30.6 and 43.9, respectively. According to the Current Population Survey November Supplements, 19.6% of eligible 18–24-year-olds reported voting in the 2010 midterm election, and 37.9% in the 2012 presidential election.

TABLE 4 CACE of Voting: Downstream Effects over Two Years

Presidential on Midterm										
	1992–94		1996–98		2000–02		2004–06		2008–10	
Arkansas			0.179	(0.058)	0.133	(0.049)	0.157	(0.019)	0.106	(0.015)
Colorado	0.156	(0.031)	0.037	(0.039)	0.086	(0.024)	0.148	(0.011)		
Connecticut					0.117	(0.026)	0.113	(0.012)	0.083	(0.009)
Iowa							0.142	(0.049)	0.038	(0.019)
Illinois	0.100	(0.019)	0.159	(0.023)	0.120	(0.014)	0.097	(0.008)	0.096	(0.006)
Florida	0.184	(0.024)	0.102	(0.019)	0.164	(0.017)	0.223	(0.013)	0.082	(0.005)
Florida 2005	0.155	(0.021)	0.060	(0.017)	0.063	(0.018)				
Kentucky									0.117	(0.012)
Michigan			0.465	(0.043)	0.129	(0.012)				
Missouri	0.215	(0.037)	0.098	(0.027)	0.110	(0.015)	0.095	(0.012)	0.090	(0.009)
Missouri 2005	0.216	(0.034)	0.082	(0.025)	0.102	(0.013)				
Montana					0.138	(0.053)	0.205	(0.038)	0.122	(0.017)
New Jersey							0.081	(0.007)	0.085	(0.005)
Nevada					0.238	(0.051)	0.150	(0.019)	0.169	(0.015)
New York							0.075	(0.006)	0.060	(0.005)
Oklahoma					0.156	(0.036)	0.112	(0.017)	0.114	(0.012)
Oregon									0.102	(0.016)
Pennsylvania							0.253	(0.016)	0.132	(0.006)
Rhode Island									0.066	(0.017)
Meta-Analysis	0.147	(0.013)	0.138	(0.011)	0.127	(0.006)	0.109	(0.003)	0.090	(0.002)

Midterm on Presidential										
	1994–96		1998–2000		2002–04		2006–08		2010–12	
Arkansas			1.122	(0.211)	0.326	(0.150)	0.330	(0.117)	0.200	(0.123)
Colorado	−0.011	(0.107)	0.074	(0.066)	0.143	(0.097)	0.088	(0.079)		
Connecticut					0.090	(0.156)	0.231	(0.092)	0.097	(0.063)
Iowa							0.228	(0.067)	0.086	(0.066)
Illinois	0.099	(0.082)	0.337	(0.092)	0.185	(0.071)	0.238	(0.067)	0.041	(0.059)
Florida	−0.047	(0.104)	0.220	(0.119)	0.064	(0.066)	0.024	(0.038)	0.014	(0.062)
Florida 2005	−0.030	(0.120)	0.131	(0.121)	0.141	(0.115)				
Kentucky									0.235	(0.052)
Michigan			0.227	(0.058)						
Missouri	0.162	(0.085)	0.260	(0.177)	0.242	(0.075)	0.129	(0.045)	0.025	(0.069)
Missouri 2005	0.243	(0.074)	0.319	(0.136)	0.101	(0.054)				
Montana					0.425	(0.099)	0.188	(0.059)	0.165	(0.091)
New Jersey							0.281	(0.092)	−0.138	(0.118)
Nevada					0.383	(0.122)	0.567	(0.129)	0.287	(0.088)
New York					0.130	(0.079)	0.156	(0.065)	0.036	(0.096)
Oklahoma					0.395	(0.158)	0.229	(0.116)	0.149	(0.090)
Oregon							0.258	(0.060)	0.118	(0.053)
Pennsylvania							0.226	(0.102)	0.104	(0.128)
Rhode Island									0.239	(0.117)
Meta-Analysis	0.068	(0.046)	0.226	(0.036)	0.200	(0.029)	0.166	(0.018)	0.111	(0.019)

Notes: Robust standard errors are in parentheses. Meta-analysis estimates exclude results from the historical voter files, Florida 2005 and Missouri 2005.

TABLE 5 CACE of Voting: Downstream Effects over Four Years

Presidential on Presidential										
	1992–96		1996–2000		2000–04		2004–08		2008–12	
Arkansas			0.246	(0.096)	0.239	(0.089)	0.221	(0.043)	0.200	(0.027)
Colorado	0.091	(0.063)	0.158	(0.057)	0.143	(0.050)	0.091	(0.025)		
Connecticut					0.289	(0.074)	0.161	(0.029)	0.161	(0.015)
Iowa							−0.024	(0.095)	0.081	(0.035)
Illinois	0.187	(0.035)	0.144	(0.043)	0.125	(0.030)	0.081	(0.020)	0.080	(0.013)
Florida	0.240	(0.045)	0.189	(0.039)	0.250	(0.038)	0.081	(0.020)	0.105	(0.012)
Florida 2005	0.253	(0.051)	0.111	(0.040)	−0.070	(0.070)				
Kentucky									0.075	(0.022)
Michigan			0.316	(0.073)						
Missouri	0.324	(0.058)	0.209	(0.063)	0.168	(0.031)	0.079	(0.022)	0.155	(0.017)
Missouri 2005	0.312	(0.052)	0.179	(0.057)	0.111	(0.022)				
Montana					0.241	(0.084)	0.222	(0.062)	0.111	(0.029)
New Jersey							0.124	(0.021)	0.155	(0.014)
Nevada					0.407	(0.093)	0.116	(0.037)	0.174	(0.027)
New York							0.132	(0.013)	0.068	(0.013)
Oklahoma					0.075	(0.076)	0.164	(0.038)	0.138	(0.023)
Oregon									0.108	(0.025)
Pennsylvania							0.203	(0.050)	0.121	(0.019)
Rhode Island									0.113	(0.030)
Meta-Analysis	0.210	(0.023)	0.189	(0.022)	0.181	(0.016)	0.117	(0.007)	0.117	(0.005)

Midterm on Midterm										
	1994–98		1998–2002		2002–06		2006–10			
Arkansas			0.466	(0.197)	0.147	(0.100)	0.212	(0.065)		
Colorado	−0.029	(0.102)	0.106	(0.053)	0.202	(0.066)				
Connecticut					0.235	(0.101)	0.209	(0.046)		
Iowa							0.097	(0.045)		
Illinois	0.168	(0.073)	0.327	(0.066)	0.178	(0.040)	0.183	(0.036)		
Florida	0.066	(0.073)	0.335	(0.098)	0.244	(0.070)	0.051	(0.019)		
Florida 2005	0.074	(0.071)	0.258	(0.089)						
Kentucky										
Michigan			0.185	(0.038)						
Missouri	0.052	(0.061)	0.387	(0.123)	0.286	(0.054)	0.060	(0.023)		
Missouri 2005	0.101	(0.067)	0.375	(0.099)						
Montana					0.272	(0.104)	0.086	(0.039)		
New Jersey							0.259	(0.038)		
Nevada					0.238	(0.075)	0.168	(0.078)		
New York					0.288	(0.036)	0.183	(0.025)		
Oklahoma					0.078	(0.093)	0.107	(0.067)		
Oregon							0.203	(0.040)		
Pennsylvania							0.073	(0.041)		
Rhode Island										
Meta-Analysis	0.075	(0.037)	0.213	(0.026)	0.232	(0.019)	0.119	(0.009)		

Notes: Robust standard errors are in parentheses. Meta-analysis estimates exclude results from the historical voter files, Florida 2005 and Missouri 2005.

in 2010 on turnout in 2012 is 0.11, with a standard error of 0.019. Looking four years earlier, the average effect of turnout in 2006 on turnout in 2008 is 0.17, with a standard error of 0.018. Voter file age is again a strong predictor of the size of the estimated CACEs, suggesting that residentially stable voters are more prone to acquire and express voting habits. Even when we look at results from a young voter file, the effects are highly significant and substantively large. Taking Oregon as our example, a CACE of 0.118 implies that casting a vote in 2010 increases the probability that a complier will cast a vote in the 2012 presidential election from 57.6% to 69.4%.

Turning next to Table 5, we track voters over four-year periods. Interestingly, the estimated four-year effects are similar in magnitude to the two-year effects. The estimated effect of voting in 2008 on voting in 2012 is 0.12 (SE = 0.005). This estimate is noteworthy for two reasons. First, this weighted average across 15 states is roughly 10 standard errors larger than the four-year persistence effect reported by Meredith (2009) based on a single pair of California elections (7 percentage points). Second, this estimate is almost identical to the effect of 2008 on 2010, which anticipates the findings described below suggesting the long-term persistence of voting habits.⁸

The apparent effect of midterm on midterm is less precisely estimated due to the smaller number of midterm compliers, but four-year persistence is nevertheless considerable. The absolute size of the coefficients is quite similar to the average midterm-on-presidential estimate, but the low base rate of voting in midterm elections means that the midterm-on-midterm effect is much larger in percentage terms.

Persistence in Turnout

An open question in the study of voter habit concerns how long-lasting the effects may be. Gerber, Green, and Shachar (2003, 549) hypothesize that the habit effect might exhibit geometric decay. Davenport et al. (2010, 429) speculate that habits might dissipate as a result of abstentions from low-salience elections. One might imagine that the effects of an eligibility discontinuity would fade over time as the political experience of the just-ineligible catches up to that of the just-eligible. On the other hand, one might imagine that habit effects accumulate as voting begets voting begets voting.

Relying on evidence from the eligibility discontinuities, we demonstrate that voting habits persist over periods of at least 20 years. Table 6 shows the CACE of voting in each upstream election from 1992 through 2010 on downstream voting in 2012. Of the 86 coefficients reported, only three are estimated to be negative (none of which reaches statistical significance). We can conduct a nonparametric test of the probability of seeing such an extreme distribution of coefficients if in fact habit effects did not persist at all with a binomial test of 83 successes in 86 trials ($p < .001$).

We present a statistical summary of the persistence of habit effects in Table 7, which pools all 384 estimates of every upstream election on every downstream election. The dependent variable in this meta-analysis is the estimated CACE, and observations are weighted by the inverse of their squared standard errors (Borenstein et al. 2009, 65). The focus of the analysis is the effect of the regressor *Years between upstream and downstream*, which refers to the length of time that elapses between the upstream and downstream election. A negative coefficient would imply that habits diminish over time, as the effects of voting when just-eligible eventually wear off. A positive coefficient would imply that habit effects become accentuated over time. A coefficient of zero would imply that the increase in vote propensity associated with voting in one's first election persists unabated over time. In order to estimate over-time decay while holding the subject pool constant, we include fixed effects for each state.

The results in Table 7 suggest little to no decrease in estimated CACEs over time. The precise magnitude of these estimates varies depending on the specification. In the first column of Table 7, we control only for the turnout rate among young voters (ages 18–29) in each state's upstream election. Column 2 adds an indicator for observations in which a downstream election took place in a presidential battleground, defined as a state in which the downstream election was decided by fewer than 10 points. Column 3 expands the definition of battleground to include closely contested downstream midterm elections for senate or governor. Column 4 introduces state fixed effects, so that the *Youth Turnout* variable represents the deviation from each state's average turnout among 18–29-year-olds, as determined by the Current Population Survey. The final column adds indicators for whether the upstream and downstream elections were presidential elections. The last model is the most predictive of outcomes, as judged by the R^2 , and generates an estimated *Years between upstream and downstream* effect of -0.0008 , with a standard error of 0.0028. This estimate implies that net of state fixed effects and election type, CACEs associated with voting habits grow weaker

⁸See the supporting information for an exploration of the robustness of these results to changes in specification.

TABLE 6 Effect of Each Upstream Election on 2012 Downstream Voting

	1992–2012		1994–2012		1996–2012		1998–2012		2000–12	
Arkansas					0.009	(0.185)	1.204	(0.451)	0.214	(0.122)
Connecticut									0.106	(0.095)
Iowa										
Illinois	0.536	(0.107)	0.494	(0.252)	0.414	(0.088)	0.968	(0.204)	0.254	(0.047)
Florida	0.688	(0.165)	−0.293	(0.376)	0.175	(0.094)	1.088	(0.311)	0.176	(0.061)
Kentucky										
Missouri	0.857	(0.177)	0.284	(0.254)	0.245	(0.129)	0.830	(0.302)	0.117	(0.039)
Montana									0.343	(0.123)
New Jersey										
Nevada									0.457	(0.143)
New York										
Oklahoma									0.195	(0.109)
Oregon										
Pennsylvania										
Rhode Island										
Meta-Analysis	0.638	(0.080)	0.264	(0.162)	0.266	(0.055)	0.986	(0.141)	0.186	(0.024)
	2002–12		2004–12		2006–12		2008–12		2010–12	
Arkansas	0.480	(0.195)	0.147	(0.045)	0.198	(0.127)	0.200	(0.027)	0.200	(0.123)
Connecticut	0.608	(0.192)	0.100	(0.027)	0.176	(0.080)	0.161	(0.015)	0.097	(0.063)
Iowa			−0.028	(0.115)	0.258	(0.078)	0.081	(0.035)	0.086	(0.066)
Illinois	0.325	(0.100)	0.150	(0.022)	0.309	(0.075)	0.080	(0.013)	0.041	(0.059)
Florida	0.046	(0.096)	0.060	(0.023)	0.060	(0.043)	0.105	(0.012)	0.014	(0.062)
Kentucky							0.075	(0.022)	0.235	(0.052)
Missouri	0.238	(0.095)	0.048	(0.023)	0.132	(0.044)	0.155	(0.017)	0.025	(0.069)
Montana	0.514	(0.153)	0.196	(0.060)	0.166	(0.060)	0.111	(0.029)	0.165	(0.091)
New Jersey			0.063	(0.021)	0.268	(0.090)	0.155	(0.014)	−0.138	(0.118)
Nevada	0.614	(0.188)	0.132	(0.042)	0.543	(0.131)	0.174	(0.027)	0.287	(0.088)
New York	0.208	(0.073)	0.071	(0.013)	0.127	(0.061)	0.068	(0.013)	0.036	(0.096)
Oklahoma	0.282	(0.210)	0.147	(0.042)	0.241	(0.120)	0.138	(0.023)	0.149	(0.090)
Oregon					0.275	(0.071)	0.108	(0.025)	0.118	(0.053)
Pennsylvania			0.243	(0.054)	0.253	(0.105)	0.121	(0.019)	0.104	(0.128)
Rhode Island							0.113	(0.030)	0.239	(0.117)
Meta-Analysis	0.271	(0.039)	0.089	(0.007)	0.176	(0.019)	0.117	(0.005)	0.111	(0.019)

Notes: Robust standard errors are in parentheses.

by only 0.008 over the course of a decade. To summarize, we find abundant evidence of long-term persistence in voting habits among those who are induced to vote upon turning 18.

Finally, we note that the composition of the subgroup to whom our estimates pertain—the compliers—changes with each election. Compliers are those who would vote in the upstream election if and only if they are eligible. Compliers who vote in low-salience elections are likely to be highly politically engaged, whereas in higher-salience elections, even relatively uninterested 18-year-olds may participate. The average vote propensity among com-

pliers is *decreasing* in upstream election salience. In an effort to model this source of heterogeneity in our treatment effect estimates, we include the youth turnout rate by state as a proxy for salience in our meta-analysis. As shown in Table 7, our estimated CACEs tend to be smaller when the overall rate of youth turnout is higher: Focusing on column 5, we see that an increase in *Youth Turnout* of 1 percentage point is associated with an average decrease of 0.16 percentage points in the \widehat{CACE} . We interpret this significant pattern as evidence that voting has a stronger habit-forming effect among higher-propensity voters.

TABLE 7 Modeling Downstream CACEs

	Dependent Variable: CACE Estimate in State-Election Pair				
	(1)	(2)	(3)	(4)	(5)
Years between upstream and downstream	0.0001 (0.0036)	-0.0006 (0.0037)	0.0002 (0.0034)	0.0014 (0.0027)	-0.0008 (0.0028)
Youth turnout in upstream election	-0.2520 (0.0635)	-0.2450 (0.0613)	-0.2578 (0.0669)	-0.2261 (0.0653)	-0.1647 (0.0779)
Presidential battleground		0.0322 (0.0147)			
Presidential or midterm battleground			0.0155 (0.0125)		
Presidential upstream					-0.0193 (0.0207)
Presidential downstream					0.0359 (0.0119)
Constant	0.2203 (0.0358)	0.2171 (0.0346)	0.2164 (0.0354)	0.2207 (0.0376)	0.2161 (0.0334)
State fixed effects	No	No	No	Yes	Yes
N	384	384	384	384	384
R ²	0.0851	0.1013	0.0998	0.2630	0.3086

Notes: Youth are defined as 18–29-year-olds. Battleground status: Two-party vote share difference is fewer than 10 points in downstream election. All models weighted by inverse of squared standard error of CACE estimate. Robust standard errors are in parentheses.

Is Persistence in Turnout Due to Environmental Influences?

The previous sections marshaled a wealth of evidence showing that random or quasi-random inducements to vote in one election also increase the probability of voting in subsequent elections. Given our maintained assumption that the inducement to vote per se has no enduring effect, we interpret this pattern to mean that the act of voting in the initial election exerts a causal effect on turnout in subsequent elections. As Green and Shachar (2000) and Gerber, Green, and Shachar (2003) point out, however, alternative causal pathways may lead from upstream turnout to downstream turnout. These pathways may be grouped into two broad categories, psychological and environmental. Psychological explanations suggest that people who are induced to vote acquire new attitudes, information, and tastes that in turn promote turnout in future elections. For example, voting may cause a person to take a heightened interest in politics, feel a stronger sense of civic duty, or become more familiar with the process by which votes are cast. Environmental explanations, by contrast, focus on mobilization activities that may be triggered by participation in the upstream election. For example, if campaigns were especially likely to

mobilize those who voted in the upstream election, we could observe that those encouraged to vote upstream are also more likely to vote downstream. This question has rarely been investigated using experimental data, but two recent attempts (Bedolla and Michelson 2012; Gerber, Huber, and Washington 2010) found little apparent relationship. Both psychological and environmental explanations are sufficient to explain the pattern we observe, but environmental explanations run counter to the notion that persistence in turnout reflects internalized habit, or what Green and Shachar (2000) call “consuetude.” Rather, environmental explanations imply that the downstream turnout effects of voting in an initial election reflect the downstream encouragements that voters subsequently receive.

In order to assess whether voting-induced mobilization activities account for the pattern we observe in our experimental and quasi-experimental data, we consider two suggestive pieces of evidence. The first comes from a closer look at the regression discontinuity results. If unmeasured mobilization efforts drive persistence in turnout, we should observe a marked difference between so-called “battleground” states, where the overwhelming preponderance of presidential campaign money is spent, and states with lopsided partisan majorities, which attract negligible presidential campaign activity. In fact, we see

little difference. The top panel of Table 5, for example, reports 15 estimates of the presidential-on-presidential downstream effect for 2008–12. The top four of the five strongest estimates (Arkansas, Nevada, Connecticut, Missouri, and New Jersey) are found in nonbattleground states.

We conduct formal tests of treatment effect heterogeneity by battleground status in columns 2 and 3 of Table 7. In those specifications, the battleground status of the downstream state election is coded as follows: In presidential years, a state is coded as a battleground if the difference in the two-party vote share in the presidential election is 10 percentage points or fewer. A similar calculation is conducted for midterm years using the gubernatorial election. If there is no gubernatorial election, we use the senate race. We code the 13 observations that did not feature races for governor or senator as nonbattlegrounds. The coefficient on the presidential battleground indicator is small, but positive and significant, suggesting that indeed, CACEs appear to be marginally stronger when the downstream election occurs in a close presidential contest. When we include the close midterm elections in the definition of battleground, this pattern does not persist—the coefficient drops by a factor of three and is no longer significant. Battleground status explains only a small portion of the observed variability in \widehat{CACEs} .

A second piece of evidence comes from a recent field experiment reported in Rogers et al. (2014). In that experiment, a group allied with the Democratic candidate for governor sent the “Neighbors” social pressure mailer in the run-up to the June 2012 recall election in Wisconsin. Rogers et al. obtained data on Democratic campaign contacts in the 2012 presidential campaign from Catalist, enabling a direct test of the hypothesis that voting in an upstream election affects campaign contact in a downstream election. They estimate that the treatment caused an 8.1 percentage point increase in the probability of receiving mail ($SE = 0.2$), a 0.3 percentage point increase in the probability of receiving a phone call ($SE = 0.2$), and a 0.4 percentage point decrease in the probability of being canvassed ($SE = 0.2$). Voting increases subsequent impersonal contacts (standard mailers), but not the personal contacts (phone calls, canvassing) that have been shown to affect turnout (Green, McGrath, and Aronow 2013). Dinas (2012, 451) finds similarly small effects of voting on subsequent campaign contact: Just-eligibles in 1968 were 12 percentage points ($SE = 8.6$) less likely to report contact in 1973, 0.5 percentage points ($SE = 10.7$) less likely in 1982, and 3.7 percentage points ($SE = 9.2$) more likely in 1997.

Taken together, the evidence⁹ from the analysis of CACEs by battleground state status and downstream effects of GOTV social pressure mailings leads us to infer that although voting affects subsequent campaign contacts, the net effect of these environmental influences appears to be small.

Discussion

The study of voting habits brings to bear an array of methodological challenges. In order to make a convincing case for the hypothesis that voting per se affects turnout in subsequent elections, the researcher must propose an identification strategy that leverages experimental interventions or as-if random discontinuities. These interventions and discontinuities must be sampled in a way that does not stack the deck in favor of the treatment group, and the immediate effects must be strong enough to avoid the statistical complications that come with weak instrumental variables. The analysis of discontinuities is complicated by asymmetries in information about the population sizes of the just-eligible and the just-ineligible. Moreover, if the downstream effect is to be attributed to habit—an acquired taste for familiar activities—rather than environmental factors or memories of the initial inducement to vote, a further layer of assumptions is required.

Given these challenges, our approach has been cautious, making use of two distinct identification strategies and an extensive set of replications. The overall pattern we observe is unlikely to be due to the idiosyncratic nature of the elections or interventions we happened to study. Although it is possible that being confronted with one’s own vote history creates indelible memories that induce turnout, the patterns we observe are too complex to be ascribed to the ongoing effects of social pressure. Participants in the August 2006 experiment were initially impelled to vote in all kinds of elections, but thereafter turnout was elevated specifically in August primaries. Given the low salience of these August elections, it seems unlikely that campaign activity targeting those who voted in 2006 would account for the elevated rates of turnout in subsequent elections. Similar arguments may be made in defense of the habit interpretation as applied to the discontinuity results. Although it is true that

⁹In the supporting information, we also report an analysis of ANES survey data according to respondents’ birthdates. Subjects who were just-eligible to vote at the time of their 18th birthday were no more likely to report campaign contact than subjects who were just-ineligible.

compliers in battleground states may have been targeted for GOTV activity in downstream presidential election years, similar CACE estimates were obtained for nonbattleground states, and significant effects are obtained in battleground states for downstream midterm elections, which are often uncompetitive. Whether one ascribes a habit interpretation to the apparent downstream effects of voter mobilization or age eligibility, evidence of downstream effects deserves the status of a stubborn fact that requires theoretical explanation.

One of the contributions of this article is to show that the pattern of downstream effects is more complex than suggested by prior work in this area, which has tended to estimate the persistent effects of voting for a single cohort and context. A more synoptic appraisal suggests that the estimated effects of habit vary significantly more than would be expected if habits took root in the same way regardless of individual differences or electoral context. The experimental results presented above, for example, suggest that downstream effects are more likely to persist in “like” elections—the inducement to vote in an August primary led to enduring effects in subsequent August primaries, less so in other types of elections.¹⁰ See the supporting information for further regression discontinuity evidence on this point.

Another benefit of assembling a broad array of statistical evidence is the opportunity to detect theoretically informative differences in the degree of downstream persistence. Results from the three large experiments suggest that downstream effects gradually dissipate over time, whereas the discontinuity results indicate that effects persist unabated even over several election cycles. These contrasting results arguably suggest individual heterogeneity in habit formation. The compliers in the discontinuity studies are 18-year-olds, who may be at a stage in life that is especially prone to habit formation. Although the hypothesis that downstream effects are larger among young adults than their older counterparts has not to our knowledge been tested,¹¹ it is consistent with evidence

suggesting that women who were ineligible to vote at the time they reached adulthood (e.g., women who reached the age of 21 in the United States prior to 1920) voted at lower rates long after they became eligible to vote (Firebaugh and Chen 1995; Merriam and Gosnell 1924).

An intriguing possibility is that both hypotheses are true: Habit formation may vary by election context and individual attributes. If so, the study of downstream effects going forward must become more ambitious in scope. Do the same discontinuity effects observed for 18-year-olds also apply to other polities, which may have different age requirements for voting? How do these eligibility effects interact with the electoral context in places such as India or Lesotho, where certain randomly selected jurisdictions are required to elect legislators from a pool consisting of only women candidates? As for experiments, to what extent do interventions that increase the motivation to vote have different downstream effects from interventions that remove or reduce administrative barriers, such as registration? Do the habit-forming effects of voting extend to other forms of political participation, such as volunteering or contributing to campaigns?

In sum, a vast body of evidence now suggests that habits form when people vote. One important implication of this fact is that events, institutions, or campaigns that mobilize voters have long-lasting consequences. In our data, we find the average CACE across all general election types to be approximately 0.10, which suggests that, *ceteris paribus*, mobilizing 100 compliers today generates 50 more votes over the five federal elections in the decade to come. Conversely, factors that demobilize the electorate have enduring repercussions, a causal relationship that may help explain why generations (Lyons and Alexander 2000) and polities (Franklin 2004) have distinctive participation profiles.

References

- Abrajano, Marisa, and Costas Panagopoulos. 2011. “Does Language Matter? The Impact of Spanish versus English-Language GOTV Efforts on Latino Turnout.” *American Politics Research* 39(4): 643–63.
- 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): 444–55.
- Aronow, Peter M., and Donald P. Green. 2013. “Sharp Bounds for Complier Average Potential Outcomes in Experiments
- text messaging (Dale and Strauss 2009) and follow-up phone calls (Michelson, García Bedolla, and McConnell 2009).

¹⁰This result is not in tension with the Franklin and Hobolt (2011) finding that those who come of age in a low-salience election context later vote at lower rates; rather, our finding states that those who are induced to vote in a low-salience election are especially likely to vote subsequently as a result.

¹¹One would ordinarily address this question by examining the downstream effects among young voters in our three experiments, but it turns out that among young people, the immediate effects of the mailings on voting are too weak to support a meaningful downstream investigation due to inadequate numbers of compliers. An intervention other than social pressure mailings may be required to experimentally evaluate the hypothesis that young people are especially susceptible to habit formation. Possible interventions that seem to have large mobilizing effects on young people include

- with Noncompliance and Incomplete Reporting.” *Statistics and Probability Letters* 83(3): 677–79.
- Atkinson, Matthew D., and Anthony Fowler. 2014. “Social Capital and Voter Turnout: Evidence from Saint’s Day Fiestas in Mexico.” *British Journal of Political Science* 44(1): 59.
- Bedolla, Lisa García, and Melissa R. Michelson. 2012. *Mobilizing Inclusion: Transforming the Electorate through Get-Out-the-Vote Campaigns*. New Haven, CT: Yale University Press.
- Borenstein, Michael, Larry V. Hedges, Julian P. T. Higgins, and Hannah R. Rothstein. 2009. *Introduction to Meta-Analysis*. Hoboken, NJ: John Wiley & Sons.
- Brody, Richard A., and Paul M. Sniderman. 1977. “From Life Space to Polling Place: The Relevance of Personal Concerns for Voting Behavior.” *British Journal of Political Science* 7(3): 337–60.
- Cutts, David, Edward Fieldhouse, and Peter John. 2009. “Is Voting Habit Forming? The Longitudinal Impact of a GOTV Campaign in the UK.” *Journal of Elections, Public Opinion and Parties* 19(3): 251–63.
- Dale, Allison, and Aaron Strauss. 2009. “Don’t Forget to Vote: Text Message Reminders as a Mobilization Tool.” *American Journal of Political Science* 53(4): 787–804.
- Davenport, Tiffany C. 2010. “Public Accountability and Political Participation: Effects of a Face-to-Face Feedback Intervention on Voter Turnout of Public Housing Residents.” *Political Behavior* 32(3): 337–68.
- Davenport, Tiffany C., Alan S. Gerber, Donald P. Green, Christopher W. Larimer, Christopher B. Mann, and Costas Panagopoulos. 2010. “The Enduring Effects of Social Pressure: Tracking Campaign Experiments Over a Series of Elections.” *Political Behavior* 32(3): 423–30.
- Denny, Kevin, and Orla Doyle. 2009. “Does Voting History Matter? Analysing Persistence in Turnout.” *American Journal of Political Science* 53(1): 17–35.
- Dinas, Elias. 2012. “The Formation of Voting Habits.” *Journal of Elections, Public Opinion and Parties* 22(4): 431–56.
- Firebaugh, Glenn, and Kevin Chen. 1995. “Vote Turnout of Nineteenth Amendment Women: The Enduring Effect of Disenfranchisement.” *American Journal of Sociology* 100(4): 972–96.
- Franklin, Mark N. 2004. *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies Since 1945*. New York: Cambridge University Press.
- Franklin, Mark N., and Sara B. Hobolt. 2011. “The Legacy of Lethargy: How Elections to the European Parliament Depress Turnout.” *Electoral Studies* 30(1): 67–76.
- Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W. W. Norton.
- Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2008. “Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment.” *American Political Science Review* 102(1): 33–48.
- Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2010. “An Experiment Testing the Relative Effectiveness of Encouraging Voter Participation by Inducing Feelings of Pride or Shame.” *Political Behavior* 32(3): 409–22.
- Gerber, Alan S., Donald P. Green, and Ron Shachar. 2003. “Voting May Be Habit-Forming: Evidence from a Randomized Field Experiment.” *American Journal of Political Science* 47(3): 540–50.
- Gerber, Alan S., Gregory A. Huber, Daniel R. Biggers, and David J. Hendry. 2014. “Ballot Secrecy Concerns and Voter Mobilization: New Experimental Evidence about Message Source, Context, and the Duration of Mobilization Effects.” *American Politics Research* 42(5): 896–923.
- Gerber, Alan S., Gregory A. Huber, and Ebonya Washington. 2010. “Party Affiliation, Partisanship, and Political Beliefs: A Field Experiment.” *American Political Science Review* 104(4): 720–44.
- Green, Donald P., and Alan S. Gerber. 2002. “The Downstream Benefits of Experimentation.” *Political Analysis* 10(4): 394–402.
- Green, Donald P., Mary C. McGrath, and Peter M. Aronow. 2013. “Field Experiments and the Study of Voter Turnout.” *Journal of Elections, Public Opinion and Parties* 23(1): 27–48.
- Green, Donald P., and Ron Shachar. 2000. “Habit Formation and Political Behaviour: Evidence of Consuetude in Voter Turnout.” *British Journal of Political Science* 30(4): 561–73.
- Highton, Benjamin, and Raymond E. Wolfinger. 1998. “Estimating the Effects of the National Voter Registration Act of 1993.” *Political Behavior* 20(2): 79–104.
- Hill, Seth J., and Thad Kousser. 2015. “Turning Out Unlikely Voters? A Field Experiment in the Top-Two Primary.” Unpublished Manuscript, University of California, San Diego.
- Huckfeldt, Robert, and John Sprague. 1992. “Political Parties and Electoral Mobilization: Political Structure, Social Structure, and the Party Canvass.” *American Political Science Review* 86(1): 70–86.
- Lyons, William, and Robert Alexander. 2000. “A Tale of Two Electorates: Generational Replacement and the Decline of Voting in Presidential Elections.” *Journal of Politics* 62(4): 1014–34.
- Mann, Christopher B. 2010. “Is There Backlash to Social Pressure? A Large-Scale Field Experiment on Voter Mobilization.” *Political Behavior* 32(3): 387–407.
- Matland, Richard E., and Gregg R. Murray. 2012. “An Experimental Test for Backlash Against Social Pressure Techniques Used to Mobilize Voters.” *American Politics Research* 41(3): 359–86.
- Meredith, Marc. 2009. “Persistence in Political Participation.” *Quarterly Journal of Political Science* 4(3): 187–209.
- Merriam, Charles E., and Harold F. Gosnell. 1924. *Non Voting: Causes and Methods of Control*. Chicago: University of Chicago Press.
- Michelson, Melissa R. 2003. “Dos Palos Revisited: Testing the Lasting Effects of Voter Mobilization.” Presented at the Annual Meeting of the Midwest Political Science Association.
- Michelson, Melissa R., Lisa García Bedolla, and Margaret A. McConnell. 2009. “Heeding the Call: The Effect of Targeted Two-Round Phone Banks on Voter Turnout.” *Journal of Politics* 71(4): 1549–63.
- Milbrath, Lester W. 1965. *Political Participation: How and Why Do People Get Involved in Politics?* Chicago: Rand McNally.
- Panagopoulos, Costas. 2010. “Affect, Social Pressure and Prosocial Motivation: Field Experimental Evidence of the

- Mobilizing Effects of Pride, Shame and Publicizing Voting Behavior." *Political Behavior* 32(3): 369–86.
- Plutzer, Eric. 2002. "Becoming a Habitual Voter: Inertia, Resources, and Growth in Young Adulthood." *American Political Science Review* 96(1): 41–56.
- Rogers, Todd, Donald P. Green, John Ternovski, and Carolina Ferreros-Young. 2014. "Social Pressure and Voting: A Field Experiment Conducted in a High-Saliency Election." Unpublished Manuscript, Harvard University Kennedy School of Government.
- Rosenstone, Steven J., and John Mark Hansen. 1993. *Mobilization, Participation and Democracy in America*. New York: Macmillan.
- Sinclair, Betsy, Margaret McConnell, and Donald P. Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments." *American Journal of Political Science* 56(4): 1055–69.
- Verba, Sidney, and Norman H. Nie. 1972. *Participation in America*. New York: Harper & Row.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

- Identification of the CACE
- Robustness of RD Estimation
- Evidence That Primary and General Elections Have Different Downstream Consequences
- Estimating the Effect of Eligibility on Campaign Contact
- Follow-up to the 2007 GOTV Experiment