The Dynamic Election: Patterns of Early Voting Across Time, State, Party, and Age

Vivekinan Ashok, Daniel Feder, Mary McGrath, and Eitan Hersh

ABSTRACT

The nature of turnout has changed in the United States: a shift in state policies has transformed a singular Election Day into a multi-week voting period. During the 2012 election, we assembled daily snapshots of early voting records across the U.S. We observe where and when individuals with key demographic characteristics voted. By measuring the timing of voting by demographic subgroups within small geographic areas, we assess how the early voting period may differentially affect various politically relevant subsets of the electorate. We find that partisans and older voters disproportionately take advantage of early voting, and that political independents and younger individuals who vote early do so much later in the early-voting window. We discuss policy implications, and we also conduct an exploratory analysis of the relationship between early vote timing and campaign events.

INTRODUCTION

Not long ago, U.S. elections were characterized by a dynamic, drawn-out campaign followed by a single, national day of voting. Over the last decade, the U.S. has experienced a “quiet revolution” in its voting procedures (Gronke 2013): the national vote is now dynamic as well. In key presidential swing states like Colorado and North Carolina, the majority of votes are cast prior to Election Day. In the country as a whole, between a quarter and a third of votes are cast early.

In this article, we study patterns of early voting across time, states, party, and age in order to shed new light on the conditions under which the electorate makes use of early voting. In the fall of 2012, we collected day-by-day snapshots of public voting records in almost every state that permits early voting. Election offices across the country provide regular reports that indicate which voters have cast early ballots. Catalist, a voter database vendor, collects and standardizes data from official reports for its subscribers. We downloaded Catalist’s early voting records in every available state, every day that such records were released. We collected counts of voters casting ballots by date, congressional district, and by key demographic characteristics. The dataset that we compiled allows us to evaluate patterns of voting behavior both cross-sectionally and dynamically. These data also allow us to overcome limitations of past studies of early voting that relied on survey self-reports, ecological data, or data from just one or two states.

Our brief analysis focuses on two key demographic characteristics—age and party affiliation—to examine behavior within politically relevant divisions of the electorate. We focus on age and party for three interconnected reasons: a) these characteristics are known to be strong predictors of voting

Vivekinan Ashok and Mary McGrath are PhD candidates in the Department of Political Science at Yale University in New Haven, Connecticut. Daniel Feder is a director at Benenson Strategy Group in New York, New York. Eitan D. Hersh is an assistant professor in the Department of Political Science at Yale University in New Haven, Connecticut.

We thank the Institution for Social and Policy Studies for research support. For helpful comments, we thank Chris Mann, Barry Burden, Paul Gronke, Marc Meredith, Deborah Beim, and participants in workshops at Yale University, Boston University, and MIT, where early drafts of this article were delivered.
behavior; b) they are recorded in the public record; and c) largely due to these first two points, electoral campaigns pay a great deal of attention to these characteristics in their mobilization efforts (Hersh 2015). Partitioning the electorate by age and party affiliation creates cogent subgroups whose early voting patterns are likely to be informative and politically relevant. In our analysis, we often distinguish between groups with higher expected vote propensities—older voters and party registrants—and those with lower expected vote propensities—younger voters and political independents. We will refer to these groups as “high-participation” types and “low-participation” types.

Our key findings are as follows: First, we demonstrate that the early voting electorate in 2012 was vastly different from the Election Day electorate. The demographic subgroups typically thought of as “high-participation” types were much more likely than “low-participation” types to cast ballots early rather than on Election Day. Second, we explore the timing of early voting ballots. We find that low-participation voters cast ballots closer to Election Day, whereas high-participation voters cast ballots closer to the start of the early voting calendar. Finally, we look at timing of the vote with respect to two types of campaign events: nationally televised debates and presidential candidate visits to a district. Though our observational evidence offers little causal leverage, we find no initial evidence that voters respond to such campaign stimuli by casting an early ballot; no consistent spikes in voting are apparent in the wake of these events among any of the subgroups we analyze. As described below, our measure of time is imprecise, and as such we may be unable to detect nuanced effects.

Our examination of early voting makes two contributions to the study of politics. First, our data allow us to examine turnout behavior in ways that were not previously possible. This dataset of day-by-day vote counts at the district level allows us to see a national picture of dynamic voting in fine-grained detail. Our over-time records provide a view of who votes when, allowing us to assess how voters might respond to key events that occur during the campaign season.

Our study also contributes to policy debates regarding the shift away from a single, national day of voting. Recent research has suggested that early voting makes elections “a more private and less intense experience,” and dampens participation among those who would have been inspired by the excitement of a single Election Day (Burden et al. 2014). By measuring the timing of voting by demographic subgroup, we weigh in on the debate about how the early voting period may differentially affect various politically relevant subsets of the electorate.

**CATALIST DATA: EARLY VOTERS FROM THE CAMPAIGN’S-EYE-VIEW**

Catalist, LLC is a political data firm that contracts with Democratic and progressive organizations as well as with academic institutions. Most of Catalist’s clients are candidate campaigns and organizations like labor unions and interest groups. These clients use Catalist’s continually updated national voter file for voter outreach. During the election season, Catalist’s clients have an interest in real-time knowledge of which voters have cast early ballots so that they can adjust their contacting strategies. Catalist obtains this information through regular data requests to local and state election officials. In most early voting states in 2012, Catalist received daily updates about early voting turnout.¹² Catalist then links reports from election officials with its own database of individual voters.

A growing body of scholarship has utilized Catalist data to measure mass political behavior in ways not possible through other data sources (e.g., Hersh and Nall, n.d.; Hersh 2013; Fraga 2016). We apply this resource to our investigation of early voting. Throughout the early voting period, we downloaded from Catalist daily counts of early voters by congressional district, party affiliation, and age. While Catalist retains past records of which voters cast early ballots versus Election Day ballots, the company did not make available the date on which the ballot was recorded as having been cast. By downloading Catalist’s records every day, we were able to gather

---

¹Catalist retrieved data less frequently than daily in Alaska, Maine, Wisconsin, and Wyoming.
²According to the National Conference of State Legislatures (NCSL), all states except the following have some combination of no-excuse absentee voting, early voting, and all-mail voting: Alabama, Connecticut, Delaware, Kentucky, Massachusetts, Michigan, Minnesota, Missouri, Mississippi, New Hampshire, New York, Pennsylvania, Rhode Island, South Carolina, and Virginia (see <www.ncsl.org/research/elections-and-campaigns/absentee-and-early-voting.aspx>). Most of our analysis focuses on turnout patterns within early voting states.
continual snapshots of the voter rolls and capture how early voting unfolds in real-time, in much the same way as campaigns observe such data.

We study simple, individual-level differences to classify voters. We distinguish registered party affiliates from unaffiliated voters (whom we will call independents), and we distinguish older voters from younger voters. Age and partisanship are important correlates of participation both to campaigns engaged in mobilization and to scholars of voting behavior. Age and party affiliation are, of course, not the only important predictors of turnout, but they are among the best predictors for which individual-level data are widely available in public records, surpassed only by measures of lagged turnout.

For example, consider the relationship between party, age, and turnout in the 2012 Cooperative Congressional Election Study (CCES). In this nationally representative survey, 49% of voters under 30 were validated voters (N=8,644) in the 2012 election. This compares with a 65% voting rate among 30–60 year olds (N=23,174) and a 75% turnout rate among voters over 60 (N=12,548). These numbers are based on all citizen respondents. However, even among registered respondents, the oldest cohort voted 16% more than the youngest cohort.

Like with age, having a party affiliation is correlated with turnout. In the 2012 CCES, 60% of self-reported independents were validated as having voted, compared to 65% of Democrats and 71% of Republicans. Among those in party registration states who are registered to vote, turnout among registrants who declined to state a party affiliation was 9% lower than for registered Democrats and 14% lower than for registered Republicans. Independents are also almost three times more likely to be undecided about whom they will vote for when asked in a pre-election survey, suggesting that they could be more susceptible to campaign interventions.

Several features of our data are particularly important to keep in mind. First, states have idiosyncratic practices with respect to early voting, in terms of the methods used to cast ballots (e.g., early in-person, early by mail), the length of the early voting period, and the interactions with other electoral rules such as same-day registration. Burden et al. (2011) demonstrate how critical it is to take account of these variations when studying the relationship between election rules and voter participation. As a result, throughout the article, we are attentive to cross-state variations.

Second, our analysis relies on a relative measure of time. The date for which Catalist records a ballot may not reflect precisely the date a voter cast the ballot because of a lag between when a ballot is cast and when it is recorded. Our analysis is sensitive to this feature of the data. Specifically, the lag between ballot cast and ballot counted may differ by state and by vote method within state (i.e., mail versus in-person). Accordingly, key parts of our analysis disaggregate data by state and ballot type.3

A third important point is that our analysis is restricted to the universe of registered voters (measured after the election) as opposed to eligible citizens. One drawback of using registered voters as the denominator is that states vary in their registration laws, creating differences in what the baseline of registered voters represents. In some states, voter registration closes before the early-voting window opens, so the pool of registered voters is identical to the pool of eligible voters. But in other states—for example, those allowing same-day registration—the pool of registered voters is a biased estimator of the complete set of eligible voters, as eligibility is not limited to those registered in advance. While recent survey-based studies of early voting (e.g., Burden et al. 2014) have been able to make use of a baseline of eligible citizens, restricting our analysis to registrants provides us the benefits of public voting records rather than self-reported turnout, along with sufficient data to study turnout by day, within demographic cohorts, within local geographies, by method of voting, and across dozens of states. To gain these benefits, we must sacrifice a consistent baseline of eligible participants. Again, this problem is largely ameliorated by disaggregating results by state or by category of state registration law.

Except when otherwise noted, we consider early voting to include all “no-excuse” methods of casting ballots ahead of Election Day, including variations of mail voting and in-person early voting, as reviewed by Gronke et al. (2008).

Regarding the timing of ballots cast, we structure voting records as survival data. In this analytical framework, the voting period begins with the full

---

3For future research, we would recommend an audit of recording lags in different states and for different ballot types. Such an audit would not only allow scholars (and campaigns) to more precisely estimate the timing of voting, but may also provide useful data on election administration.
population of voters or eligible voters within a given subgroup, and each day we observe the number that voted and therefore exited the initial pool. Following this method’s often counterintuitive nomenclature, survival indicates individuals who have not yet voted; failure indicates voting.

Motivating hypotheses

Our objective in utilizing Catalist records to study early voting is to shed descriptive light on the early voting process and motivate future research. There are essentially two lines of scholarly research upon which we build. The first is explicitly about early voting and how much convenience it adds to the electoral process. The second line of research is focused on timing and campaign effects, an area of study that takes on a new complexity when voting, and not just campaigning, is a dynamic process.

First, given that early voting has been introduced by states explicitly to make the voting process more convenient, empirical studies of early voting have given most attention to theories based on convenience (Gronke and Toffey 2008). The most common hypotheses have focused on “high-participation” types who make up their minds early in the election season about whether they will vote (Stein 1998; Karp and Banducci 2001; Gronke and McDonald, n.d.; Gronke 2013; Barreto et al. 2006; Neeley and Richardson Jr. 2001). Once voters decide that they will vote, they choose the most convenient time to cast their ballot. Thus, individuals who are nearly certain they will vote from the beginning of the campaign season will take advantage of early voting more than those who tend to be less sure about their turnout decision. Those who make up their minds early on tend to be partisan, older, and more engaged with politics, and as such, these attributes should be correlated with early voting. These voters tend to know both that they will vote and for whom they will vote months ahead of an election. Many empirical studies of early voting are consistent with this explanation, which is why Berinsky (2005) concludes that early voting fails to reduce participatory bias in elections, a “perverse consequence” of election reform.

In our analysis, we will show in sharp detail how distinct types of voters differentially take advantage of early voting. Specifically, compared to the higher turnout cohorts of older voters and partisans, younger voters and independent voters disproportionately cast ballots on Election Day, and among those who vote early, they typically vote close to Election Day.

Turning to the second line of research, scholars like Shaw (1999), McClurg (2004), Holbrook and McClurg (2005), Hillygus (2005), and Hillygus and Jackman (2003), have studied campaign events such as debates, party conventions, and candidate appearances, examining how turnout intention and vote intention shift following such stimuli. Given the opportunities that voters have to cast ballots weeks ahead of Election Day, we can reconsider hypotheses of campaign effects not by measuring turnout intention in the midst of the campaign but by measuring actual turnout in the midst of the campaign. Voters may react to specific events—like televised presidential debates and local visits by presidential candidates—that occur during the early voting period. If this is true, we should see increases in turnout at particular times in the early voting calendar (after major events) and in particular places (locations where campaign news and events are particularly salient and where voters can vote early in-person even if they are not previously registered). Our second motivating hypothesis is that early voting rates may temporarily increase in the wake of televised debates and local candidate visits.

ANALYSIS

Early vs. Election Day voting

The bulk of the literature on early voting suggests that demographic groups with typically higher turnout—partisans and older voters—will be more likely to cast early ballots than will lower-turnout types. Figure 1 shows that the 2012 election data comport with these expectations. The plots on the left show early voting; the plots on the right show Election Day voting. The upper plots use age to differentiate high- and low-participation types, while the lower plots use party affiliation. In the upper plots, the x-axis represents the turnout rates among registrants under 30 and the y-axis represents the turnout rates of older registrants. Two different categories of “older” registrants are shown: gray triangles show the turnout rate of registrants aged 30–60, while black circles show the turnout rate of registrants over 60. In the lower plots, the x-axis represents the turnout rate
FIG. 1. Early and Election Day turnout rates of high- and low-participation types, by state. Plots compare the early turnout rate and the Election Day turnout rate among low-participation cohorts on the x-axis with high-participation voters on the y-axis. The upper plots are based on full registered population statistics, totaling 103,339,432 people in 25 states. The lower plots are based on 67,494,299 records in 17 party-registration states.

among registrants with no party affiliation (party-registration states only) and the y-axis shows the turnout rate among registered Democrats (gray triangles) and registered Republicans (black circles). A triangle or circle falling on the dashed 45-degree line indicates that the turnout rates among high-participation cohorts and low-participation cohorts in that state are the same.

The upper-left plot of Figure 1 demonstrates that the early voting electorate skews toward older voters. This “participatory bias” toward older voters is especially pronounced in the states where early voting is most popular. In states where 50%–80% of older registrants are utilizing early voting, only about 20%–40% of younger registrants are voting early. Shifting to the upper-right plot, the Election Day electorate is shown to be considerably different. In half the states, a greater proportion of registered younger voters are casting ballots on Election Day than older voters. However, the heteroscedasticity in the left plot is not mirrored in the right plot, implying that states that have especially high rates of early voting among older voters are not also showing especially higher rates of Election Day voting among younger voters. The lower plots, which measure turnout rates by party cohort, exhibit a similar pattern as
the age plots. Partisans of both parties are more likely to cast early ballots; independents are more likely to vote on Election Day.

Figure 1 shows that high-participation types have higher turnout rates in the early voting period than low-participation types. The differences across left and right plots in Figure 1 are important for understanding contemporary U.S. elections. For example, the patterns suggest a new complexity in assessing data obtained through Election Day exit polls, one of the most common sources of data measuring mass behavior (see McDonald and Thornburg 2012). More immediately, the differences exhibited are consistent with convenience-based theories of voting in which citizens who are generally more certain to vote are also more likely to take advantage of early voting. The difference in early voting and Election Day turnout rates visible in Figure 1 also contradicts single-state, survey-based research (e.g., Barreto et al. 2006) that has found minimal differences with respect to political party affiliation between early voters and Election Day voters.

**Early-vote timing and participatory bias over time**

We now evaluate the evidence for two expectations that would comport with previous research: first, that high-participation types are more likely to cast their ballots at the start of the early voting period, and second, that as early voting progresses and as media attention to the election builds, low-participation types vote at increasingly higher rates.

In Figure 2, we plot Kaplan Meier failure estimates for age and party cohorts. These graphs show cumulative turnout patterns for each cohort. The steps in the plot indicate when new data appeared in the Catalyst database. For each point in time, the y-axis indicates the proportion who have voted up to that point. The left plots are conditional on having voted early, whereas the right plots are conditional on registration. The right plots end at the proportion of registrants within the cohort who voted at any point in the election. For example, in the upper right plot, by the end of the early voting period, just under 40% of registrants over 60 had already cast a ballot, and just over 60% of this cohort had cast a ballot by the end of Election Day. The left plot ends at 1.00 on the y-axis because by the end of the early voting period, all early voters have cast a ballot.

Figure 2 summarizes data only for states that have relatively long early voting periods. For inclusion in this graph, we required a state to have an early voting period beginning at least two weeks ahead of Election Day. This excludes Florida, which had an unusual and truncated early voting calendar in 2012 (Herron and Smith 2012; Gronke and Stewart III, 2013). We also exclude Oregon and New Mexico, for which we do not have data on Election Day voting. Note that while we include all citizens who are subset in alternative ways, such as focusing just on same-day-registration states, or just on presidential swing states, or just on the subset of votes that are cast early in-person or early by mail. For one illustration, in Figure A1 in Appendix A, we break out the left plots of Figure 2 by method of early voting.

Consider the left-side plots of Figure 2, showing cumulative turnout among early voters. In both the age and party plots, the low-participation voters and high-participation voters exhibit a distinct vertical gap from approximately October 24th (two weeks ahead of Election Day) to November 1. Among the early voters, this means that the high-participation groups are voting earlier than the low-participation groups. Among those who cast an early ballot, nearly 50% of younger voters cast their early votes after November 1; fewer than 25% of older voters cast a ballot that late in the early voting window. The pattern is similar, though less stark, for the party cohorts. Independents, like the younger voters, were more likely to cast their early ballots very close to Election Day. Registered Democrats appear to vote slightly earlier than registered Republicans, but both groups of partisans vote earlier than independents.

The right-side plots, showing cumulative turnout among registrants, reveal a turnout gap between high-participation voters and low-participation voters that grows larger over time. On October 24th, 14% of registered voters over 60 had already cast ballots, compared to 3% of voters under 30, an 11 percentage point gap. By the end of the early voting period, 36% of older voters had cast ballots compared to 16% of younger voters, a 20 percentage point turnout gap. A similar pattern is apparent with party cohorts. The increasing gap in participation between low-participation and high-participation voters is not a foregone conclusion. Since low-participation early voters are more likely to vote closer to Election Day, if more registrants in this cohort opted to vote early, the gap between high-participation and low-participation voters could narrow at the end of the
early voting period. But because early voting is more popular among high-participation voters at every point in the calendar, the participatory gap, or participatory bias, increases with time.

Compared to the high-participation voters, early voting rates among young and independent registrants start low and stay low throughout the early voting period. But low-participation types who do decide to cast an early ballot are much more likely than high-participation voters to cast their ballot in the final days before an election. Perhaps for these voters it takes more time to gain enough information, willpower, or interest to cast an early ballot.

The most important policy implication of this analysis is that lengthening the early voting period on the front end by opening it earlier may not be a particularly helpful reform for low-participation voters. Low-participation voters, such as political independents and young voters, are most likely to vote last-minute, either on Election Day itself or in the days immediately preceding Election Day. Reforms that are intended to start the early voting period earlier are therefore more likely to add convenience for the high-participation voters who will vote in any case. On the other hand, extending or reducing the number of early voting days closer to Election

FIG. 2. Change in participatory bias over time, by party and by age group. The y-axis indicates the proportion voting by a particular date. Only states that have an early voting period lasting at least two weeks are included. In the upper plots, included states are AK, AZ, CO, GA, IA, ID, KS, ME, MT, NC, ND, NE, NV, OH, TN, TX, UT, WA, WI, and WY. In the lower plots, included states are AK, AZ, CO, IA, ID, KS, ME, NC, NE, NJ, NV, UT, and WY.
Day is more likely to impact low-participation voters. For those reformers interested in expanding voting opportunities for low-participation voters, it appears more important to ensure that the days immediately preceding Election Day are open to voting as opposed to days much earlier in the calendar.

**Turnout response to campaign events**

Figure 1 shows that the use of early voting is concentrated among traditionally higher-participation cohorts of older voters and partisans. We now look for evidence suggesting that campaign events may create deviations from this regularity. Voters whose turnout decision is impulsive or spontaneous may be especially likely to cast an early vote after key events, even if their overall turnout rates are not catching up to those of high-participation types over time. For citizens on the “turnout bubble,” early voting provides a long window of opportunity to spontaneously cast a ballot. We ask whether low-participation voters have noticeable spikes in turnout in the days following these events.

To evaluate this event-based hypothesis, we focus primarily on Iowa and Ohio. These are not only presidential swing states, they are also same-day registration states (Burden et al. 2014), which means that eligible citizens could decide to spontaneously cast a ballot even if they were not previously registered. These are also states with both early in-person and early mail voting. Because one need not request a ballot prior to voting in-person, in-person voting may be better suited for studying spontaneous mobilization in response to events. These states also have a very long window of early voting, lasting more than a month. In studying responsiveness to national debates, we thus focus on Iowa and Ohio. In studying local visits by presidential candidates, we focus just on Iowa. Iowa is a state where congressional districts are compact and clearly distinguish four regions of the state. Because our data are stored geographically by congressional district, Iowa’s clean district lines mean that it is straightforward to determine which votes in our dataset were close to a presidential candidate visit.

In Figure 3, we plot the timing of statewide early voting for Ohio and Iowa. We show age data for Ohio and party data for Iowa. Ohio does not have party registration, and Iowa’s age data are consistent with the results displayed for Ohio. The left plots show early in-person voting; the right plots show early by-mail voting. Since presidential debates are some of the biggest national events during a campaign season, and since all the debates took place within the window of early voting in these states, we seek evidence that turnout spiked following these events, especially among low-participation voters.

Notice in Figure 3 that early in-person voting is more prevalent later in the campaign season than mail voting. For some cohorts, nearly half of the mail ballots cast were processed before October 15th, whereas a majority of in-person voters did not cast ballots until the very last days of the campaign. As a result of this, the first three debates took place before noticeable numbers of in-person voters were casting ballots.

To the extent that there are any bumps in turnout following presidential debates, they are clearly larger among high-participation cohorts than among low-participation ones. This is most visible among the mail voters. For example, consider the upper-right plot, from Ohio data. A few days after the second and third presidential debates there is a spike in turnout, but that spike is much larger for the oldest cohort than for the youngest cohort. Of course, just observing turnout over time does not provide much causal leverage to make claims about voters responding to events. But, at least in our viewing of these data, there are no especially compelling signs low-participation voters are stimulated by debates to spontaneously cast a ballot. The turnout increases following the events for low-participation voters are similar to or smaller than turnout increases for high-participation voters.

Our exploration of turnout responsiveness to debates does not indicate that low-participation voters were especially reactive to these national events. But perhaps these voters are responsive to local events like visits by presidential candidates. In Figure 4, we identify the dates of visits by presidential candidates Romney and Obama and the congressional district

---

4 Note in observing the plots in Figure 3 that while Catalist received daily updates from election authorities, and while we pulled data from Catalist on a daily basis, not every day indicates noticeable amounts of voting.
in which the visits took place. Each sub-plot in Figure 4 shows when early voters are recorded as having voted in Iowa. Republican candidate Romney visited each of the four districts once. President Obama visited the first district twice, and two other districts once. Note that because all of these events happened in late October and early November, for the sake of visual clarity Figure 4 shows data from October 15th to November 6th.

There are two noteworthy features of Figure 4. First, while there are some increases in turnout following local candidate visits to Iowa districts, those bumps have inconsistent lags following the visits. This leads us to suspect that the bumps in turnout may not be connected to the campaign events. This suspicion might be wrong, especially if different parts of the state have different lags between when ballots are cast and when they are recorded. Second, compared to the partisan cohorts, the day-by-day turnout rate of independent voters is unexceptional. The turnout increases are not noticeably

---

5The smallest geographic unit of analysis for which we have data is congressional district, which is why we use this geography rather than other plausible geographic indicators.
larger among independents than among partisans, nor are they larger among younger voters than among older voters (not pictured).

Our examination of voter responsiveness to major campaign events is necessarily blunt: we measure turnout before and after the event, and we see no obvious signs of increased turnout among low-participation types or high-participation types following the event. We looked for day-to-day surges of voting where we were most likely to witness some movement consistent with a mobilization story: swing states with same-day-registration and long windows of in-person early voting. We see no strong evidence that voters respond to campaign events by casting an early ballot. To be clear, voters may be mobilized by such events in ways we do not see, but the detailed dataset that we have obtained from Catalist does not show clear evidence of a mobilization story.

**DISCUSSION**

Using the largest database of early voting turnout ever analyzed in political science research, we have learned a number of lessons about early voting. First, confirming the longstanding view that high-participation voters are most likely to take advantage of early voting, we showed that in most early

---

**FIG. 4.** Iowa early turnout by date, in relation to presidential candidate visits. Plots show timing of vote recorded by Catalist among all Iowa voters who voted early in-person, by congressional district. Observations counts, in order of district number, are 55,026; 79,386; 57,761; and 61,007.
voting states, older and partisan registrants, who tend to be reliable voters overall, were indeed more likely to cast early ballots in 2012. This means that the Election Day electorate tends to be disproportionately composed of low-participation voters. Indeed, on Election Day itself, it is often the case that independents are voting more than partisans and that young voters are voting more than older votes.

Second, we found that among those who do cast an early ballot, high-participation groups are voting earlier in the pre-election window than are the low-participation groups. That is, when younger voters or those not registered with a party vote early, they do so much closer to Election Day than do older voters and party registrants. This finding is important for policymakers who are deciding on the duration and placement of the early voting window: lengthening or shortening the early voting calendar at different ends of the voting period would have distinct impacts on different kinds of voters.

Finally, we sought evidence for whether voters respond to news-making campaign events, like nationally televised debates and local visits by presidential candidates, by casting ballots in the days following these events. We see no clear evidence suggesting that voters respond to such campaign-season stimuli by casting early ballots. While bumps in turnout are visible, we find nothing large or systematic enough for even suggestive causal attribution. However, our data are observational and the measure of timing is noisy; more precise measurement and experimental strategies could reveal effects that remained undetected in the present analysis.

This study constitutes the largest-scale descriptive account of early voting yet undertaken, but it has some important limitations. For one, our pool of abstainers is constrained to registrants listed in the Catalist database. Particularly in states that allow for same-day registration, use of the registered population as the denominator is a biased estimator of the complete set of eligible voters. While we acknowledge this limitation, it is worthwhile to keep two points in mind. First, our findings do not change when we restrict the states under investigation to only those that permit same-day registration. And second, we believe that the biases from use of registration records instead of a measure of citizen voting age population would mute, rather than bolster, our key findings. For example, consider that young people are not only less likely to vote conditional on registration; they are also less likely to register conditional on being eligible (Ansolabehere and Hersh 2013). Because the pool of unregistered young people is proportionately much larger than the pool of unregistered older people, if anything our findings are likely to understated the large degree of participatory bias in voting. While use of registration records as a denominator for eligible voters may be imperfect, it represents a trade-off with considerable up-sides. Namely, the use of the Catalist records rather than survey data provides us the benefits of public voting records rather than self-reported turnout and also sufficient data to study turnout by day, within demographic cohorts, within local geographies, by method of voting, and across dozens of states.

A second limitation of this study is our choice to focus on just two key explanatory variables: age and party. These are two of the most important correlates of voting in American politics, but we do not suggest that they tell us the whole story about who votes when and why. Downloading full registration updates every day in all of these states was an extraordinarily labor-intensive process, and we needed to make choices about what variables to capture; vote method, local jurisdiction, age, and party were our highest priorities. This leaves room for future research to expand upon our design. With many universities now subscribing to Catalist data, and with official data more widely accessible than ever before, we are hopeful that others will build on our efforts.

The contribution of this study is in evaluating real-time snapshots of individual-level early voting behavior. Our observational approach has all the familiar limitations regarding causal inference, and our evidence regarding turnout in the wake of campaign-season events (debates and candidate visits) should not be interpreted as estimates of causal effects. Our hope is that the descriptive evidence we provide will prompt future research in these areas. The newly dynamic nature of U.S. elections means that the nature of voter turnout has changed dramatically. The dynamic election leads to a far more complex process of voting that political

---

6For a future study, we would recommend obtaining more precise estimates of the timing of the vote, perhaps by supplementing Catalist records with state records. We would also recommend downloading data on each voter’s history of prior turnout in addition to voters’ demographics, geography, and partisanship.
scientists are just beginning to understand. By studying patterns of voting across time, states, party, and age, we can gain insight into the contours of the new electorate.

REFERENCES


Address correspondence to:
Eitan Hersh
Department of Political Science
Yale University
77 Prospect Street
New Haven, CT 06520
E-mail: eitan.hersh@yale.edu

(Appendix follows →)
APPENDIX A

Notes on state categorizations

According to the National Conference of State Legislatures (NCSL), all states except the following have some combination of no-excuse absentee voting, early voting, and all-mail voting: Alabama, Connecticut, Delaware, Kentucky, Massachusetts, Michigan, Minnesota, Missouri, Mississippi, New Hampshire, New York, Pennsylvania, Rhode Island, South Carolina, and Virginia.

The following 27 states (and D.C.) allow no-excuse absentee voting: Alaska, Arizona, California, Colorado, D.C., Florida, Georgia, Hawaii, Idaho, Illinois, Iowa, Kansas, Maine, Maryland, Montana, Nebraska, Nevada, New Jersey, New Mexico, North Carolina, North Dakota, Ohio, Oklahoma, South Dakota, Utah, Vermont, Wisconsin, and Wyoming.


The following 10 states (and D.C.) allow same (election) day registration: California, Connecticut, D.C., Idaho, Iowa, Maine, Minnesota, Montana, New Hampshire, Wisconsin, and Wyoming.

Some of our analysis distinguishes swing states, also known as battleground states, from safe states. Nine states in 2012 were widely considered battleground states. These states are Colorado, Florida, Idaho, Illinois, Iowa, Nevada, New Hampshire, North Carolina, Ohio, Virginia, and Wisconsin. Of these states, all but New Hampshire and Virginia are early voting states.

Notes on state data

Since all voting in Oregon is considered “early” (though voters can return ballots on Election Day, these ballots are not separately identifiable in the dataset) and some of our analysis compares early voting to Election Day voting, Oregon is excluded from those parts of the analysis. For Washington State, which, like Oregon, has all-mail voting, Catalist is able to distinguish mail ballots that were submitted in-person on Election Day from ballots sent by mail ahead of Election Day. Accordingly, we are able to compare Election Day voting to early voting in Washington. Note also that we were unable to obtain Election Day turnout figures appended to Catalist records in New Mexico for 2012 by the time our contract with Catalist expired. Several early voting states are not party registration states, so they are excluded from analyses that compare party affiliates to independents.

Several states have particularly noteworthy idiosyncrasies. First, Illinois is the only early voting state that did not provide Catalist with pre-election updates of early voting turnout. Illinois is thus the only early voting state that is entirely excluded from our analysis due to data limitations. Second, Michigan presents an unusual case. Michigan is not considered an early voting state because voters need an excuse to request an absentee ballot. However, one of the sanctioned excuses in Michigan is being at least 60 years old. Many older voters take advantage of absentee voting in Michigan: Catalist recorded over one million early votes cast in that state, nearly all of which were associated with older voters. Because early voting opportunities in Michigan are constrained by age, we leave Michigan out of most of our analysis. Finally, in five states (Idaho, Kansas, Louisiana, Nebraska, and Ohio), the tallies for Election Day turnout as generated from the individual-level records maintained by Catalist appear to be the sum of Election Day voters and early voters. In these states, when we seek to compare Election Day and early turnout, we must subtract the early voting counts from the Election Day counts. The origin of this particular feature of the data seems to be that these states do not distinguish early from Election Day voters in their voter files.

It bears acknowledging that in addition to obtaining data about voters’ vote method, vote timing, geography, age, and party, we also downloaded data from Catalist’s file based upon Census block group level characteristics, such as percent of a voter’s block group with a college degree and percent non-Hispanic white. However, we found the block group level to be unreliable. The patterns of behavior suggested by the block group analysis appeared odd, and we learned from other scholars who are

using the Catalist data that for some period of time in 2012, the process of linking individuals in the Catalist file to Census characteristics may have been inaccurate. We do not have sufficient confidence in the block group data to include it in our analysis.

FIG. A1. Figure 2 (left panels only) plotted separately for in-person and mail ballots. The y-axis indicates the proportion voted by a particular date. Only states that have an early voting period lasting at least two weeks are included. In the age in-person voting plot, states include AK, IA, KS, ME, NC, NV, OH, TN, UT, and WA. In the age mail plot, states include AK, AZ, CO, GA, IA, ID, KS, ME, MT, NC, ND, NE, NV, OH, TN, TX, UT, WA, WI, and WY. In the party in-person voting plot, states include AK, IA, KS, ME, NC, NV; and UT. In the party mail voting plot, states include AK, AZ, CO, IA, ID, KS, ME, NC, NE, NJ, NY, UT, and WY.