The Consequences of Primary Election Timing[1]

Robert G. Boatright, Clark University[2]
Vincent G. Moscardelli, University of Connecticut
Clifford Vickrey, University of Connecticut

June 2017

Abstract

There has been substantial disagreement about the "best" time to hold primary elections, yet there is little evidence about the effects of different election dates. This paper represents the first attempt to utilize a newly-assembled dataset covering primary and general elections from 1984 to 2016 to assess the consequences of primary election dates for turnout, cost, and competitiveness of primary elections to the U.S. House of Representatives. Searching over this full time period, we uncovered no evidence of a substantively meaningful effect of primary timing on either the cost or competitiveness of primaries. However, turnout does appear to dip in the summer months before recovering in the fall—a pattern that has become more pronounced in both parties in recent years. Descriptive statistics and bivariate analyses suggest that in the past decade, later primaries have become more expensive; however, we found no evidence of this pattern once we controlled for potentially confounding forces. Future research will be necessary to determine whether this recent shift represents an aberration or an emerging trend in primary election dynamics. We conclude with a brief discussion of proposals regarding the optimal timing of congressional primary elections and some thoughts on the relationship between primary timing and ideological polarization.

[1] The authors gratefully acknowledge the financial support of the William and Flora Hewlett Foundation's Madison Initiative. Any errors are the responsibility of the authors. In accordance with the Hewlett Foundation's open access policy, replication data and the full dataset assembled for this project are available at https://wordpress.clarku.edu/primarytiming/. Thank you as well to Rebecca Hadik, Joya Hall, and Jana Kelnhofer for research assistance.
[2] Corresponding author (RBoatright@clarku.edu).

The Consequences of Primary Election Timing

The increase in polarization within the U.S. Congress over the past two decades has led to a renewed interest in the consequences of primary election laws. Primary elections, first introduced in the early twentieth century, were expected at that time to change legislative elections in dramatic ways, reducing the power of political parties, increasing the cost of election campaigns, and empowering voters in states where one party had traditionally been dominant. During the first decades of the twentieth century, researchers asked many questions about the effect of primaries and about the effects of different types of primary laws. However, as primary elections became an accepted feature of American politics, and as American political parties found ways to exert control over primaries, political scientists gradually lost interest. The recent ideological sorting of the parties, however, has been accompanied by increased concern by parties with shaping primary contests (Hassell 2016), increased interest group activity in primaries (Boatright, Malbin, and Glavin 2016), and increased media attention to ideologically extreme candidacies within both parties (Boatright 2013). Primaries, in short, have become more interesting than they were for much of the twentieth century.

Many of the original questions asked about primaries have by now been answered, and for the most part the answers have suggested that there are few modifications to primary laws that yield measurable consequences. Researchers have studied the effects of party endorsements in primaries (Dominguez 2011, Hassell 2013), variations in who is allowed to vote in primaries (Norrander and Stephens 2012, Gerber and Morton 1998, Kanthak and Morton 2001, McGhee et al 2014; Rogowski and Langella 2015), and different types of runoff provisions (Gerber and Morton 1998, Lamis 1984, Bullock and Smith 1990) on outcomes of interest such as turnout and competitiveness, but for the most part, results in these areas have been slight. In a handful of cases primary rules have certainly made a difference in election outcomes, but researchers have failed either to find consistent effects from any set of primary election characteristics or to corroborate normative claims about what the "best" set of primary rules would be.

In our judgment, the most consequential early question about primaries that has not been answered is the question of what the effects of primary election timing are. We have long known that political parties, in the United States and abroad, seek to manipulate election dates in order to maximize their chance of victory. It stands to reason, as well, that dominant factions or candidates within political parties will similarly manipulate election timing to advantage themselves. Such manipulations are standard practice in election systems where the governing party can call an election (Docherty 1997; Ramiro 2016) and, in some instances, in state and local politics (Anzia 2014). Yet in the American system, federal general election dates are fixed and primary election dates, while not fixed, generally are established sufficiently before the election that manipulating the date for short-term advantage tends to be more difficult.

Establishing primary election dates, therefore, requires institutional strength and a rationale that extends beyond the election at hand, and that might serve as a plausible normative story. In other words, anyone seeking to shift the date of a primary from, for instance, March to May (or May to September), would need to pass legislation to do so, and could expect some public scrutiny and skepticism of the endeavor.

There is a variety of folk theories about the normative consequences of election timing, most of which arose during the early twentieth century, and none of which have been conclusively tested. In his 1908 book *Primary Elections*, Charles Merriam (pp. 137-139) summarized the thinking of the time: if primaries are held in the spring, candidates will effectively have to wage two entirely different campaigns, and the issues of political importance may change substantially. If the primary is held during the summer, voters will be less interested; rural voters, in particular, may be busy with their crops (see also Horack 1923). If the election is held during the fall, lesser known candidates may be disadvantaged going into the general election and intraparty conflicts in the primary will not be healed. Similarly, European political scientist Ernest C. Meyer (1902, 399) opined that summer vacations would diminish primary turnout but fall primaries might prevent parties from being able to develop general election campaign strategies for their candidates. Geiser (1923) contended that early primaries would yield more costly general elections for candidates; in an era where parties tended to fund general election campaigns and campaign finance laws tended not to be applied to primaries, it appeared that primary timing could have a substantial effect on whether election campaigns were funded by parties or candidates (Baker 2012, 47).

A century later, no one has proven or disproven these claims, although many state-specific accounts have made further claims about party preferences regarding primaries. Histories of states such as Florida (Craig and Austin 2008), Virginia (Sabato 1977), and New York (Zimmerman 2008) allege that political parties have tended to dislike late primaries because there is insufficient time for intraparty conflicts to be resolved in time for the general election. Zimmerman, in fact, contends that the majority party in New York has at times adjusted the primary date for the explicit purpose of harming the minority party. On the other hand, in 1951 the National Municipal League issued a set of recommendations about helping parties assert greater control over primary elections; among these was a recommendation that states set a primary date in early October, with filing deadlines in early September. The rationale here was that early primaries made general election campaigns unnecessarily long; these lengthy general election campaigns had, it was argued, become "an unnecessary bother and expense to the candidates and to the public as well."

This confusion is reflected in the current panoply of state primary election dates (shown in Table 1). There is little logic to these dates or the changes in them. Some states with a tradition of strong political parties, such as Illinois and Ohio, have very early primaries; other strong party states, such as the New England states, are among the last states to hold primaries.[3] States with a history of weaker parties, such as the western and plains states, are similarly spread across the spring, summer, and early fall. No state holds its primary as late as the National Municipal League had suggested.

[Table 1 about here]

Since the early 1970s, there have in addition been questions raised about the wisdom and consequences of concurrent presidential and nonpresidential primary elections. Studies have shown that voting in nonpresidential primary races is higher in states where the presidential and

[3] For useful distinctions among state party systems, see Mayhew 1986, Elazar 1972.

nonpresidential primary elections are on the same ballot, particularly in early presidential primary states or in years where the primary contest is competitive (Boatright 2014, ch. 3).[4] Some previous research has controlled for the effect of the coincidence of presidential and nonpresidential primaries on the representativeness of the electorate or on the ideological characteristics of party nominees (Herrnson and Gimpel 1995; Kaufmann, Gimpel, and Hoffman 2003; Johnson, Petersheim, and Wasson 2010). Anecdotal accounts have often suggested, for instance, that liberal candidates and African-American candidates for lower office benefitted from appearing on the same primary ballot as Barack Obama in the 2008 primaries (Helderman 2008). This adds an extra wrinkle to analyses of primary timing – concurrent primaries tend to occur earlier than nonconcurrent ones, but variations in the characteristics of the electorate or the campaign are likely determined not by the election date but by the presence of a presidential primary on the ballot (and perhaps by the characteristics of the presidential candidates).

Although most of the early twentieth century literature discussed above is largely speculative or impressionistic, several recent studies have considered primary timing in the context of larger models of primary election turnout or competition, but results have been mixed. Johnson, Petersheim, and Wasson (2010; see also Lazarus 2005) investigate the relationship between primary election timing and general election divisiveness, concluding that early, divisive primaries can be harmful to incumbents. Galderisi and Ezra (2001), in contrast, contend that late primaries are more divisive for the party, while also observing that states with traditionally strong political parties tend to hold their primaries earlier. Kaufmann, Gimpel, and Hoffman (2003) contend that at the presidential level, early primaries can advantage more ideologically extreme candidates. Two dissertations on primary election rules (Sabella 2009, Kurlowski 2014) use primary election timing as a variable in models predicting turnout or competition; both conclude that it has little effect. Another recent dissertation uses primary dates to determine whether members of Congress change their voting habits or other legislative activities at the time of their primaries (Schmitt 2013). These studies vary in the ways in which they conceptualize election timing; some measure the length of the general election campaign (e.g. the time between the primary and the general election) while others distinguish between early, middle, and late primaries. In a study of Spanish primary elections, Ramiro (2016) contends that early primaries can boost a party's general election vote share while later primaries reduce it. However, the effects he uncovers are nonlinear; there is a tipping point for Spanish primaries of about eight months before election. While there is no reason to expect U.S. elections to exhibit similar temporal features, it is possible that the way in which primaries dates are conceptualized and operationalized – e.g., as a simple count of the number of days or months between the primary and the general election, a non-linear transformation of one of those counts, or as separate categorical variables for early (spring), middle (summer), or late (fall) – might be consequential. In fact, in the analysis that follows, we demonstrate that analysts' conclusions

[4] Between 1984-2016, contested incumbent primary turnout was two percentage points higher in races held on the same day as a presidential primary (12.8 percent) than in races that did not occur on the same day as a presidential primary (10.7 percent). Turnout in challenger primaries, which exhibit lower turnout than incumbent primaries to begin with, witnessed a similar reduction (6.4 percent when held concurrently; 5.7 percent otherwise). Both of these differences are statistically significant at $p \leq .05$ level. Turnout in open seat primaries was not as high as in incumbent primaries, and surprisingly, we uncovered only a substantively negligible difference in turnout rates in open seat primaries that were held the same day as a presidential primary (10.1 percent) and those held on different days (9.5 percent).

about the relationship between timing and turnout in primary elections may be sensitive to how *primary timing* is operationalized.

Modern literature on primary timing, then, is sparse and inconclusive, despite the various folk theorems about what its effects should be. This is the case in part because of difficulties measuring it, and in part because it has rarely been a direct object of study.

Literature on primaries tends to focus upon four different subjects. First, there are arguments about voter turnout and the composition of the electorate. In these studies, claims are made about whether more people are inclined to vote at different times of year, or how the composition of the primary electorate might be affected by spring, summer, or fall primaries. Second, claims tend to be made about competitiveness in both the primary and the general election. As regards the primary, it has been argued that early primaries tend to advantage incumbents and to discourage competition, in part because voters may not be as engaged in the electoral process as they will become as the general election approaches. As to the general election, prominent arguments concern the effect of primary divisiveness on general election outcomes, and the amount of time it might take to calm feathers ruffled by a divisive primary. Third, varying claims have been made about whether early primaries are more expensive than later primaries, or whether early primaries can make general election campaigns more expensive. To some extent, it seems evident that the earlier the primary, the more time there will be for primary victors to raise money for the general election. But it has also been argued that some types of spenders (parties or access-oriented PACs, for instance) will be advantaged or disadvantaged according to the timing of the primary. And finally, arguments have been made about whether or not nominees with certain qualities – e.g., party regulars or insurgents, ideological moderates or extremists — will be advantaged or disadvantaged by different primary dates.

These four areas of inquiry overlap, to be certain. Primary competition, for instance, may be a consequence of turnout, spending, or the qualities of the candidates. And, as noted above in the discussion of concurrent primaries, many individual primary elections do not by themselves determine spending or turnout. House primaries, in particular, may be shaped not only by presidential races but by Senate or gubernatorial primaries. Any effort to isolate the effects of primary timing on outcomes of interest must control for a variety of other state, national, or candidate-specific factors that may swamp the effects of the primary date. To the extent that we can isolate timing effects, however, we are able to present a clear story. When we consider the entire 1984-2016 period, timing does not have a major effect on either spending or competition in primary elections. Turnout, however, does appear to dip during the summer months—a trend that has become more pronounced in both parties in recent years. Beyond this, our conclusions are more tentative. Utilizing descriptive/univariate and bivariate techniques, we present evidence that Democratic primaries have become less competitive and more expensive as the general election draws closer in this decade, suggesting an effort by the Democratic Party to reduce competition. During that same period, similar analyses reveal that Republican primaries have exhibited higher-than-expected turnout and competition late in the primary cycle. These differences, we speculate, may corroborate recent arguments about political polarization and differences in party culture. However, these shifts, which we observe in a small sub-sample of our most recent data, are modest in magnitude and tend to vanish once we control for potentially

confounding forces. Thus, future research must revisit these questions before analysts can be certain whether or not the patterns we uncover during the 2010s constitute an historical aberration or an emerging trend in the dynamics of primary elections.

Research Questions, Data, and Evidence

For the purposes of this paper, we focus our analysis on the relationship between our key independent variable – the timing of primaries – and three dependent variables of interest: turnout, cost, and competitiveness of U.S. House primary elections. We are unable to measure directly the effects of primaries on candidate ideology, although we do speculate at the end of the paper about these relationships.

In keeping with previous work, we conduct separate analyses for incumbent primaries, challenger primaries, and primaries for open seats. This paper represents our initial attempt to use a newly-constructed dataset to answer a basic question: does the timing a congressional primary affect the cost, turnout, or competitiveness of primary elections to the U.S. House? The data on which we rely are aggregated at the level of the primary election and cover the 1984-2016 period, so the unit of analysis is the primary-district-year.[5] That is, for each congressional district in each election year, there may be two cases – one for the Democratic primary and one for the Republican primary. Our dataset consists of three types of variables: variables relating to the election itself, variables relating to election outcomes, and variables relating to the financing of the election. In addition, we collected data on a variety of different candidate, district, and state characteristics from a variety of different sources. A fuller summary of information on the construction of this dataset, with emphasis on the spending variables, is available in Appendix A. Here, we briefly summarize the most pertinent details on the data we collected, with an emphasis on the variables we use in the analyses that follow.

Dependent Variables

Turnout: For years 1994 and later, federal election returns for primaries were taken from the Federal Election Commission's (FEC) election results webpage. For years 1992 and earlier, as well as for gubernatorial candidates, research assistants entered raw vote totals from printed matter (Scammon and McGillivray, various years) into a cloud-computing-based web form.[6]

[5] All told, we assembled substantial amounts of information on the more than 51,000 candidates who received votes in primary elections between 1970-2016. Our campaign finance data on primary elections go back to 1980. See Appendix A for a discussion.

[6] After they were collected, the House primary vote totals were cross-checked with returns collected by James Snyder, which allowed us to triangulate and reconcile any inconsistencies that emerged, thus further reducing errors in our own data. Calculating turnout for uncontested races can present problems because raw vote totals are not reported in every district in every year. Therefore, below we restrict our analysis to contested primaries and general elections – operationalized here as those in which more than one candidate received more than a handful of votes. We will explore the effects of alternative operationalizations of "contested" primary and general elections in future iterations. However, for most (though not all) of the bivariate analyses that appear below, we also restricted the

General election returns were taken from the FEC, as well as the Inter-university Consortium for Political and Social Research's (ICPSR) historical election series. House general election returns were checked via comparison with Gary Jacobson's historical House of Representatives spending data. Senate and gubernatorial returns were checked by hand. District-level estimates of voting age population (and total population) were taken from various Census Bureau sources. These allow us to compute not only the number of people who voted in each party primary, but to calculate turnout as the percentage of the VAP who participated in any given primary election. Turnout in primary elections tends to be quite low, with an average of about nine percent of the VAP participating in primaries across all states, all primary types, and both parties during the period studied here.

Candidate Expenditures: Cycle-wide campaign finance totals are readily obtainable. The FEC has made total receipts and total disbursements available in its summary files back to the 1980 election cycle. However, a student of Congressional primaries runs into a simple problem with cycle-wide finance figures: the totals of primary winners, who must raise and spend money after the primary election, are incomparable with those of primary losers. The mean(median) contested House primary from 1970 to 2016 took place 137(154) days before the general election, so that eventual nominees on average campaigned for almost five months longer than their defeated opponents. If our research question is (say) whether primary timing and ballot access deadlines affect campaign finance, this discrepancy will defeat any attempt at an empirical test. Unfortunately, the FEC's online campaign summary records furnish no easy way to aggregate pre-primary transactions. The process we used to acquire these data is somewhat complicated; this, as well, is described in Appendix A. But in a nutshell, we used date-stamped pre-primary and pre-convention transaction records to assign all receipts and disbursements to either the primary or the general election period. We then used this information to generate aggregate totals for receipts and disbursements in each primary-district-year.

Our major financial variable of interest here is expenditures, not receipts. While many researchers (including us; see Boatright 2013, ch. 4) focus their attention on candidate receipts, in this instance we use expenditures as a means of looking at campaign activity rather than candidate strength. That is, a strong candidate may raise a substantial amount of money, deter potential opponents, and ultimately spend little of it during the primary, while another may raise a similar amount of money and need to spend all of it in order to win a close election. One can ask questions about both activities, but if we are interested in the effects of primary dates, our questions have to do with the nature of the primary—specifically its cost—not the nature of the candidates.

We collected other financial variables as well, including total PAC contributions, contributions from PACs of different types, party contributions, unitemized contributions, and candidate loans and contributions. All are of interest in general elections and all undoubtedly play some role in primaries. However, the vast majority of primary candidates raise so little money that we do not provide extensive analysis of these variables here.

---

cases to those primaries in which no single candidate received greater than 95 percent of the vote and the results were not affected.

Competitiveness: We measure competitiveness differently for the three primary election types. In the case of open seat and challenger primaries, we follow several extant studies (Canon 1978, Herrnson and Gimpel 1995, Hogan 2003, Brogan and Mendilow 2012) in employing a fractionalization index which is calculated as

$$F = 1 - \sum [(C_1)^2 + (C_2)^2 + (C_3)^2 + (C_4)^2 \ldots]$$

where F is the fractionalization index, C_1 is the percentage of the total vote received by the first candidate, C_2 is the percentage of the total vote received by the second candidate, and so on. This yields an index where a one candidate race has a fractionalization index of zero and a race where two candidates split the vote would have a fractionalization index of 0.5 (or $1 - (0.5^2 + 0.5^2)$). The larger the number of similarly competitive candidates, the closer the index is to 1 – that is, a race with ten candidates who received ten percent of the vote each would have an index of $1 - [(.1)^2 * 10]$, or 0.9. The intuition behind these indices, in other words, is that an election where one candidate gets most of the votes is not very fractionalized, even if there are multiple candidates; races with two candidates with similar vote share are split, and those with more than two equally competitive candidates are even more divided. High values on this variable are associated with more "competitive" races.

The fractionalization index is adept at capturing differences in competition in races where competition between multiple candidates is the norm. For incumbent primaries, however, we would contend that fractionalization is not always the best indicator of competition because, as noted above, the vast majority of incumbents run without serious primary opposition. Thus, following Boatright (2013) we distinguish here between incumbents who run without a serious opponent and incumbents who were held to 75 percent or less of the primary vote.[7] We thus have a binary measure – either incumbents faced a credible challenge or they did not. We also calculated the incumbent's margin of victory (or defeat) over the second place finisher in the election being considered.

Independent Variable of Interest: Election Dates and Timing

Our independent variable of theoretical interest is the timing of the primary election. Using FEC documents, we recorded the date of every House, Senate, and Presidential primary election in every congressional district from 1970-2016. We also recorded the candidate filing deadlines for these party primaries and general elections dating back to 1984. We used a variety of print and electronic sources to identify the dates of gubernatorial primary elections and recorded those as well. When do House primaries occur? Figure 1 is a histogram of the distribution of House primaries over the course of the year in election years from 1980 and 2016. Primary elections tend to be clustered in one of three "seasons" – with over half of all primaries falling in what could be categorized as the spring (62.7 percent), and a smaller number falling in summer (21.1 percent) and fall (16.2 percent).[8] During this period, the median contested primary

[7] When we restrict the analysis in this way, approximately one in four incumbent primaries meets the threshold.
[8] Primaries categorized as "Fall" generally took place roughly in September and early October. "Summer" primaries were held from roughly the middle of June through the end of August. "Spring" primaries fell before mid-June. No

fell 154 days before the general election – a number that remains stable across the entire time series. The mean, on the other hand, is 143 days, and ranges from a low of just 121 days in 1976 to 171 days in 2000. The large spike in the summer reflects the fact that some large states like California, Ohio, and New Jersey all generally held their primaries exactly 120 days before the general election for much of the time period covered here.

[Figure 1 about here]

In the discussion of results, we look at primary timing both as a continuous variable (that is, number of days between primary and general), as one of three seasonal categories, and as one of eight month categories (February-September). There is no one correct way to do this. However, we would contend that distinguishing between seasons has three benefits: first, many of the folk theories about primaries have had to do with seasons – when people tend to take vacations, when crops are harvested, and so forth. Thus, changing the primary date from one season to the next may have an effect that changing the date from, for instance, March to April will not. Second, distinguishing between seasons also removes the noise introduced by small legislative decisions or by natural rhythms of the calendar (for instance, a primary mandated to be held on the second Tuesday in May will vary slightly in its proximity to Election Day from one year to the next). And third, it is simply easier to look at differences across time periods when one displays them by season, especially when the changes in the dependent variables are nonlinear. After we use the seasonal categories to explore and present broad patterns in the data, we then employ the monthly categories in the multivariate analyses that follow. As we demonstrate below, the choice of operationalization has no discernible effect on our results when it comes to cost and competitiveness of primary elections. However, in the case of turnout, results are sensitive to the choice of operationalization of primary timing.

Control Variables

We also collected data on a large number of potential control variables for use in this and other studies. In the various analyses that follow, we also employ a number of control variables to capture other sources of variation in primary turnout, spending, and competitiveness. These include:

- Characteristics specific to the primary itself, including the party holding the primary; whether the primary is held concurrently with the presidential primary or not; whether the primary his held concurrently with a senate or gubernatorial primary; whether the concurrently held primary is competitive; and the voting rules for the primary (that is, whether it is open, semi-open, semi-closed, or closed). We exclude nonpartisan "top two" or blanket primaries from our analysis here, although the dataset we created does include those races.

- Characteristics specific to the political context. These include year and state fixed effects, a variable for the decade, and district Democratic presidential vote share in the

traditional (non-blanket; non-special election) primary fell fewer than 30 days before the general election; and the earliest primary in our data fell 273 days before the election.

current or previous presidential election. We also use two commonly accepted measures of state political culture, David Mayhew's total party organization (TPO) measure of party strength (1986) and Daniel Elazar's (1972) measure of state political culture. We use the ICPSR region categories to distinguish between areas of the country.

- Characteristics specific to the candidates. Our dataset includes measures of traits and past activities of incumbents, including NOMINATE scores, age, seniority, and past vote share. We also include some measures for nonincumbents, such as the Bonica (2014) CF scores of candidate ideology, and measures of candidate quality, prior political experience, and candidate occupation.[9]

Many of these variables have been found to be strong predictors of electoral performance; others we have included in our dataset so that interested readers can explore potential determinants or other types of questions related to primary election activities and results.

Approach and Results

In the pages that follow, we explore the relationships described above in three ways. First, we present some basic bivariate analyses (bar charts, scatterplots and correlations). These patterns serve as the starting point for our second approach, a multivariate investigation of each of our dependent variables. Third, after presenting and discussing the aggregate results, we then turn to a more granular analysis of both intra- and inter-state variation in primary dates and their relationships to turnout and competitiveness.

Turnout in House Primaries and General Elections

We turn our attention first to the relationship between primary timing and turnout. Some studies have contended that turnout will vary in different seasons, although the literature has not necessarily suggested a linear trend. Many early studies have noted that turnout will be highest in the fall, when voters have devoted more thought to the election, and lowest in the summer, when many voters are engaged in typical summertime activities (e.g. Meyer 1902, Merriam and Overacker 1928). Some recent work on turnout that is *not* specific to primaries supports such claims (Anzia 2014). It has also been argued, variously, that late primaries can generate either momentum or divisiveness in the general election – it is not clear whether such things might inspire higher or lower general election turnout, but in both instances there is the potential that some voters will be inspired to show up (or not) in the election as a consequence of a late primary while the effects of an early one will have dissipated by November.

[9] Prior occupation data for 2000-2010 are gathered from Stephen Pettigrew's Harvard dataverse site, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/26448. Our research assistants added data from 2012-2014 to our file. More description of this process is provided in the appendix.

Figure 2a presents scatterplots, linear trends (with 95 percent confidence intervals), and Pearson's correlation coefficients that describe the relationship between our key independent variable – the number of days that passed between the primary and the general election – and turnout (measured here as percent of the district VAP).[10] Spring primaries (which take lower values on the x-axis) appear to enjoy somewhat higher turnout than summer or fall primaries across all three primary types (incumbent, challenger, and open seat). Of course, the actual turnout rate (approx. 9 percent on average across all states, all primary types, and both parties) is still quite low. Furthermore, the on-average differences are not large or statistically significant at traditionally accepted levels. For example, turnout averages 9.4 percent in spring primaries and 8.4 percent in fall primaries (a difference of 1 percent or less than 7,000 voters in the average primary). Nonetheless, the strong, statistically significant negative correlation ($r = -.22$) between primary date and turnout in incumbent primaries stands out and merits further investigation going forward. Interestingly, Figure 2b reveals that districts with late primaries do appear to experience slightly higher turnout in the general election. This is true in both open seat races and races involving incumbents. An average of 155 primary voters drop off each day that passes after the first primary of the season, and on average, districts with fall primaries have approximately 40,000 fewer voters than districts with spring primaries on general election day. The causal mechanism underpinning this relationship is not clear, but intuition points to the possibility that the primary election may indeed "prime" voters to participate in the general election in November.

[Figures 2a and 2b about here]

If we leave aside the relationship between primary and general election turnout, however, to solely look at causes of variation in primary turnout, two additional causes of seasonal variation emerge. First, as our calendar of primary dates (Table 1) shows, the majority of concurrent presidential/nonpresidential primaries happen in the spring. Second, we may be capturing differences between parties, whether because some parties' voters may have different levels of commitment to the primary or different seasonal constraints on voting, or because states where one party is dominant may have primaries at different times than those that are more competitive or where the other party is dominant. New England, for instance, tends to have later primaries than the rest of the country, and New England has over the time period considered here been more Democratic than the rest of the country.

Let us first consider concurrence. To what extent are variations in turnout across time merely capturing differences between the presidential and nonpresidential primaries? Although it might seem intuitive, it is in fact somewhat difficult to develop a method for comparing concurrent and nonconcurrent primaries. Not all presidential primaries are competitive, either because one party's presidential candidate is running unopposed, because the nomination has been effectively wrapped up by the time of a given primary, or because one candidate has enough of an advantage in a state that the primary is not really contested. In order to avoid judgment calls in this regard, Figure 3 simply compares primaries held in presidential election years to those held in midterm election years. Even with this loose level of comparison, it does

[10] So that later primaries will appear farther to the right side of the x-axis, we reversed the scale of this variable. As a result, the x-axis in all the scatterplots that follow captures the number of days that have passed between the primary in question and the first primary of that election cycle.

seem highly likely that the presence of a presidential primary on the ballot drives most of the variation in turnout. Specifically, in presidential election years, turnout drops noticeably across the seasons as some candidates drop out of the presidential race and others sew up their parties' nominations.

[Figures 3a and 3b about here]

Figures 4a and 4b show variations in turnout by party; Figure 4a shows all primaries, and Figure 4b excludes New England. Both figures show that, at this level of aggregation, turnout tends to decline in a somewhat linear fashion from spring to summer to fall. A comparison of the two figures shows, furthermore, that the New England states are not responsible for this pattern. Of the six clusters of party primary types shown here, the only one that does not decline across seasons is Democratic challenger primaries, which exhibit a dip in turnout in the summer before recovering in the fall.

[Figures 4a and 4b about here]

Another interesting characteristic of the relationship between political parties and election timing is the change in the relationship between partisanship, timing, and turnout across the years considered here. Figures 5a and 5b compare the pattern exhibited in the 1980s to that of today. During the 1980s, the steady decline in primary turnout from spring to summer to fall is evident across all primary types and within both parties. By the 2010s, overall turnout was lower (that is, if one looks at the y-axis, it is clear that average turnout was lower for most primary types by the 2010s; the decline for Democrats is particularly obvious). Yet the decline across seasons in turnout is gone by this point. During the 2010s, some late primaries have actually had higher turnout than those held earlier—a pattern never observed in the 1980s.

[Figures 5a and 5b about here]

These graphs thus tell contradictory stories about the relationship between turnout and timing. This may speak to an increasing interest on the part of voters in congressional primaries (or a declining interest in presidential primaries), and it may also suggest that some late congressional primaries have become more important to the types of voter who participate in the low-turnout affairs that 21st century congressional primaries have become.

Cost of House Primaries and General Elections

Some of the accounts of the introduction of primaries that we listed above (particularly Baker 2012) contended that the direct primary would make campaigns more costly; some arguments about timing held that early primaries would lead to more expensive elections overall (Hempstead 1901). According to this line of reasoning, candidates would effectively have to finance two entirely different campaigns if primaries were early, whereas a late primary would more easily segue into the general election campaign. Late primaries, however, may also yield longer campaigns. If, in addition, the late primary is not very competitive but the general

election campaign is expected to be, more money directed at the general election will be spent during the primary.

We uncovered no substantively meaningful relationships between levels of overall spending in primary races (of any type) or general election races and the date the primary election was held. Figure 6a plots the natural log of total spending by all candidates in House primaries from 1984-2016 against the timing of the primary, broken out by type of primaries (incumbent, challenger, and open seat). In all three cases, the bivariate analysis uncovers no substantively significant pattern to this relationship. Similarly, Figure 6b plots the natural log of total spending by all candidates in general election races between 1984-2016 against the timing of the primary, broken out by primary type. Again, the bivariate analysis reveals no pattern between the timing of the primary and total spending in the general election. Of course, total spending is not the only meaningful indicator of the role of money in congressional elections, and there are many other variables to consider – especially those involving the flow of money to candidates from parties and groups, whose practical opportunities to spend money in general election races may be constrained in states that hold later primaries.

[Figures 6a and 6b about here]

The lack of an aggregate pattern, however, masks interesting variations within different types of primaries. As in the previous section, we again consider variations between presidential and nonpresidential years, between the two parties, and across the decades considered in this study.

Figure 7 compares spending by House primary candidates running in presidential election years to spending by these candidates in midterm election years. In midterm election years, there is no general pattern to the relationship between timing and expenditures; instead, each type of race exhibits a distinct pattern. In presidential election years, however, much less money is spent by candidates running in late primaries than by those running in earlier ones, and in every case (incumbent primaries, challenger primaries, and open seat primaries), spending peaks in the summer of presidential election years. The low levels of spending observed in fall primaries that take place in presidential election years may correspond to the turnout variations we saw in the prior section, but it is less intuitive, especially since the declines we observed in turnout were monotonic, whereas spending peaks in the summer before bottoming out in the fall. It is easy to conclude that more voters show up in early primaries because there are presidential candidates on the ballot. Yet why should the presence of a presidential race on the ballot prompt House candidates to spend more money? One could hypothesize that competition from the presidential race for voters' attention prompts other candidates to spend more money in order to get voters' attention. Candidates running after the presidential nominations are decided, in contrast, may not need to work as hard.

[Figure 7 about here]

Figures 8a and 8b compare spending by Democrats and Republicans across the decades considered in this paper.[11] One must keep in mind here the change in partisan control of

[11] Note that the dollar amounts here are not adjusted for inflation.

Congress: during the 1980s, Democrats held a comfortable majority in the House, and thus there were simply more Democratic incumbent primaries than there were Republican incumbent primaries. In 2010, in contrast, control of Congress changes hands; throughout this period the margin of control on the part of Republicans is narrower than it was for Democrats during the 1980s. The ability of different types of candidates to raise and spend money, and the stakes involved in these elections, in other words, were undoubtedly different.

[Figures 8a and 8b about here]

One intriguing pattern that emerges in these figures, however, is that for the 2010s, spending tends to be somewhat higher in later primaries than in earlier ones. This is particularly the case among Democrats. Our hunch, however, is that this variation may merely capture expenditures made during the primary for the purpose of the general election. It is difficult to distinguish between spending geared toward winning a primary and spending geared toward winning a general election, especially in the case of the latest primaries in the cycle. As Boatright (2013, ch. 4) notes, it is safe to assume that a candidate who spends money in primary where he or she has no opponent is really spending that money in order to win the general election; when that candidate has a primary opponent, however, we are forced to make assumptions about the purpose of the spending using imprecise guesses based on things like the opponent's total spending, the ultimate vote share, or other imperfect measures of *a priori* competitiveness.

It is difficult to reconcile the findings regarding parties here with the findings regarding concurrent presidential election. The differences here suggest to us, however, that primary timing has no clear influence on spending. This is not to say that primary timing does not influence other campaign finance variables. Yet such variables are difficult to measure in the aggregate because they are so dependent on the competitiveness of the primary itself.

In Figures 9 and 10, however, we show two noteworthy patterns in the behavior of different types of donors. Political Action Committees (PACs) are often seen as politically savvy "investors" – that is, they will favor candidates who are likely to win. This is particularly the case for corporate PACs. Figures 9a and 9b we show the average receipts, by race, for two types of contributions, PAC contributions and corporate PAC contributions. Unsurprisingly, the vast majority of both contribution types go to incumbents; we have not separated out incumbents and their primary opponents, but one can assume (based on Boatright 2013, ch. 4) that the vast majority of these contributions are to the incumbents. When we turn our attention to primaries that consist of other candidate types, there is some evidence that later primaries draw more PAC and corporate PAC contributions, likely because this money is, in part, geared toward helping the expected primary victor prepare for the general election.

[Figures 9a and 9b about here]

Another donor category to which we might impute strategic implications is small donors – individuals who give less than $200. We can measure such contributions because candidates are not required to itemize these contributions – they must merely report them as a lump sum. Conventional wisdom holds that small donors tend to be more ideological or more prone to give to candidates who are unique or exciting. An abundance of small donations is also indicative of

grassroots campaign work by nonparty groups. Examples of candidates who have excelled at raising such small contributions include Senators Elizabeth Warren and Bernie Sanders, on the Democratic side, and Representatives Ron Paul and Michelle Bachmann on the Republican side. Figure 10a shows the mean total of unitemized receipts by primary. In contrast with PAC contributions, these small contributions tend to be distributed more uniformly across the three different types of races, with open seat primaries, on average, receiving the largest total amount in small contributions. This is to be expected, given the greater competitiveness of open seat races and the lack of the sorts of ties to PACs that incumbents would have. However, in further contrast with PAC contributions, small individual contributions tend to wane over the course of the primary season. In every type of race in both parties, primary races that take place in the fall attract less money from these sources than races that take place in either the spring or summer. This is notable because it suggests that late primaries, on average, may be less exciting to small donors.

[Figures 10a, 10b, and 10c about here]

Figures 10b and 10c compare unitemized receipts reported by Democrats and Republicans during the primary season across the decades for which these totals are available.[12] At least one noteworthy difference does emerge over time. In the 1990s, late primaries tended to attract less, or approximately the same amount, in contributions from small donors than early primaries. But by the 2010s, with the exception of Republican challenger primaries and Democratic incumbent primaries, late primaries tended to attract more – and in some cases much more – in contributions from small donors than early ones. While our data do not allow us to draw conclusions about the forces behind this shift, we speculate that at least two are at work. First, this shift may reflect changes in the information environment in which donors, candidates, and parties operate. The growth of the Internet and significant outreach and targeting efforts by parties may results in more donations being channeled toward a small number of high profile races. Second, the advent of independent groups like MoveOn.org and ActBlue, who encourage small donors to contribute to high profile, close races, may also have contributed to this trend.[13]

In sum, there is little evidence that primary timing in itself influences campaign spending or contributions; rather, most of the evidence we have presented suggests that not all primary expenditures or contributions are specifically about the primary itself. For much of the period considered here, more money was spent by congressional candidates in early primaries. However, this pattern seems to have changed recently. Late primaries are more costly not because of any seasonal dynamics but simply because the primary election season captures some campaign activity that is about the general election. Whatever hypotheses people may have had in the early twentieth century about primary election spending may have made sense at the time but are not necessarily applicable to today's elections. That said, different patterns do emerge among different types of donors. Not only are incumbent primaries characterized by heavy PAC

[12] We collected unitemized receipt data for the 1990-2016 period. As with previous figures, readers should note that the dollar amounts here are not adjusted for inflation.

[13] In fact, using inflation-adjusted figures (constant 2017 dollars), the standard deviation of mean unitemized contributions during the primary season was 1.5 times larger in the 2010s than in the 1990s. While not definitive, this is consistent with a pattern of some primary races continuing to attract very little in the way of small contributions while others began to attract very large amounts of money.

activity relative to challenger and open seat primaries, but PAC spending in incumbent races clearly peaks in the summer (this is also true when we restrict the analysis to corporate PAC contributions). In contrast, small individual donors appear to be more active open seat primaries and in spring and summer primaries. By fall, when more sophisticated actors allocating PAC funds are directing more contributions to challenger and open seat primaries (perhaps in an effort to improve the expected primary winner's position heading into the general), contributions from small individual donors have dropped off, in some cases precipitously.

Competitiveness in House Primaries and General Elections

Third, we turn our attention to competitiveness. Claims about the relationship between timing and competition are more sparse in the literature, but there are some. Most of these have to do with party control; some claims have been made that "strong" parties will hold primaries early, in order to confer the nomination on their preferred candidate before anyone is really focused on the election; others have argued just the opposite, that strong parties will hold their primaries later, in order to provide more time for the field to be winnowed (Boyd 1989; Elazar 1999). These opposing views have a "just so" quality to them, in that some states with notoriously strong parties, such as Illinois, tend to hold their primaries very early, while other "strong party" states such as New York and Massachusetts, tend to hold theirs very late. Fractionalization is not a perfect measure of whether party-preferred candidates are advantaged, but if parties have successfully winnowed the field, one would expect lower levels of fractionalization to result.

As described above, we use fractionalization to measure competitiveness in challenger and open seat primaries. Figure 11a plots fractionalization for both types of primaries against the date on which the primary was held. Challenger primaries appear to be equally competitive across the range of the electoral calendar. However, open seat primaries appear to be slightly more competitive in the spring than they are in the fall ($r = -.10$). This turns out to be different than the pattern we observe in incumbent primaries. Figure 11b plots the percent of primary elections in three "seasons" (spring, summer, and fall) deemed to have been "competitive" using two thresholds of different stringency – (a) the incumbent received less than 75 percent of the vote; and (b) the incumbent received less than 90 percent of the vote. Regardless of the threshold employed, and contrary to the patterns we observed in challenger and open seat primaries, the earliest incumbent primaries tend to be the least competitive, and primaries held in the summer tend, on average, to be the most competitive.

[Figures 11a and 11b about here]

We noted above that turnout in early primaries seems to be largely driven by concurrence with presidential primaries. There is no obvious reason to assume that fractionalization would be affected by concurrence, but, as Figure 12 shows, it does appear that challenger and open seat primaries are slightly less competitive in the fall of presidential election years than they are in the spring and summer. There is no apparent relationship between fractionalization and season during midterm election years. It is possible that the higher turnout in concurrent primaries creates some uncertainty and can benefit candidates who would have a more difficult time in

nonconcurrent races. One way to think about this is to remember that if voters are showing up because of the presidential race, they may well have lower levels of political information, or lower levels of support for the political status quo. The presence of these voters may yield more competitive primaries in part out of ignorance and in part out of the fact that these voters may simply be different in their political views.

[Figure 12 about here]

Variations by party across the seasons largely accord with the patterns we have seen so far. Democratic primaries have become less competitive over the past three decades and Republican primaries have become more competitive. Of particular note in Figures 13a and 13b are the changes in competitiveness in challenger primaries. During the 2010s, late Democratic challenger primaries are less competitive than early ones, but late Republican challenger primaries are more competitive than early ones. In the 1980s, in contrast, late primaries of all types were less competitive than early ones for Democrats, and were never *more* competitive than early ones for either party. We saw above that turnout is lower in late primaries, so this competitiveness is not necessarily related to turnout.

[Figures 13a and 13b about here]

So, there is, again, a difference between the full 1984-2016 time series and the elections of the 2010s. For the full time series, early primaries appear to be more competitive than later ones. Yet for the 2010s, late Republican primaries are just as competitive as early ones. As was the case for turnout and spending, recent primaries show different seasonal dynamics.

Overall, then, we have evidence that there is nothing structural about the primary calendar that has a major effect on turnout, spending, or competition. That is, most of the observed differences during the full 1984-2016 period seem likely to be driven by concurrence or expected general election competitiveness. It is possible to construct a multivariate model that tests for these factors, as well as for other determinants of change in our three dependent variables. However, variation across the 1984-2016 period suggests that a separate analysis of recent primaries is also in order. It is to this task that we now turn.

Multivariate Analyses

The preceding discussion explores the relationship between primary timing and our three dependent variables of interest (turnout, cost, and competitiveness) using two different operationalizations of primary timing – the number of days between the primary and the general election and the season during which the primary took place. In this section, we explore these relationships in more detail while controlling for prominent alternative explanations and/or confounding forces. We estimated all of the models presented below as cross-sectional time-series generalized least squares (GLS) regression with random effects and robust standard errors clustered by a variable that uniquely identifies the party holding the primary in a given district

(our panel identifier). As was the case in the preceding discussion, the unit of analysis remains the individual primary election. The large number of cases in our data allows us to include dummy variables for individual states and individual election years to account for effects unique to any given state or any given election cycle.[14] We also control for a variety of theoretically- and empirically-informed covariates. While we present the results of these models in tabular form for interested readers, our primary tools for interpreting our results are plots of predicted values derived from those results. Thus, the discussion that follows draws primarily on the results depicted in Figures 14, 15, and 16.

Independent Variables and Controls

The independent variables in which we are most interested center on primary timing and primary type. To assess primary timing, we depart from the continuous (days between primary and general) and seasonal (spring, summer, fall) indicators used above and rely primarily on an indicator of the month in which the primary occurred. In most of the models presented below, we then interact this month counter with an indicator of the type of primary – incumbent, open, or challenger – and/or with an indicator for the party holding the primary.[15]

We also include several controls. First, we include a dummy variable identifying states that fall in the upper two categories of Mayhew's Traditional Party Organization scale. While Mayhew's typology includes five categories, the binary categorization is appropriate for our purposes since states with high values on Mayhew's scale can be characterized by "strong" party organizations, which some have speculated might more effectively take steps to suppress primary competition. We include the natural log of total spending in the race as an indicator of the cost and prominence of the race.[16] Because it could affect the size and composition of the electorate, we include dummy variables for whether or not there was a (1) presidential or (2) gubernatorial primary held on the same day. We also control for the rules surrounding primary participation – e.g., is the primary classified as "closed," "semi-closed," "semi-open," or "open" ("open" primaries are the reference group). Finally, in all the models that follow, we include lagged values of the dependent variable.

Turnout

As described above, we calculated turnout as the percentage of the voting age population in a district that participated in an individual primary election. Table 3 and Figure 14 present the

[14] In the tables that follow, to conserve space, we do not report the estimates associated with these dummy variables. Complete results available from authors upon request.

[15] We used Stata's `xi` features to generate the interactions and Stata's `margins` and `marginsplot` routines to generate the predicted values plotted in the figures that appear below. Code for estimating these models and generating these plots will be made available on the Clark Primary Timing website.

[16] Total spending in a race can, of course, both affect and be affected by an election's competitiveness. Accounting for this recursivity empirically requires any number of solutions, including the use of two-stage models with instrumental variables. Fortunately, such techniques not necessary for our purposes as our goal is not to estimate without bias the precise magnitude of either of these effects. Instead, we simply control for the money plays in these races and focus our attention on the relationship between primary timing and our three outcomes of interest.

results of our GLS regression of turnout on our independent variables. All four models fit the data well (overall R^2 values range from .48 to .50). In the first, we include our indicator of the number of days between first primary of the cycle and the primary in question (so that high values on this variable are associated with later primaries). These results are presented in Table 2a, and the associated predicted values are plotted in Figure 14a. Figure 14a presents a negative overall relationship between timing and turnout – i.e., later primaries appear to experience lower turnout. Plotting the predicted values from this model, we see that incumbent and open seat primaries enjoy higher turnout than challenger primaries, and that later primaries experience lower turnout. However, diagnostics reveal that imposing this linear trend on the results may overstate the nature of the relationship, especially given the high turnout usually witnessed in March and the low turnout sometimes witnessed in July. Indeed, when we aggregate by month and estimate the same model (with categorical month interactions), Table 2b and Figure 14b reveal that predicted turnout in open seat and incumbent primaries hovers between low (13 percent) and lower (9 percent), with no clear over-time trend. Challenger primaries exhibit the lowest levels of turnout, but beyond a dip in July, they also reveal no meaningful pattern as the cycle progresses. Table 2c and Figure 14c present the results of model of primary turnout specified to assess the relationship between primary month and turnout, broken out by primary type and party. Again, challenger primaries exhibit a dip in turnout during July, but otherwise, the range of predicted turnout rates remains narrow, and the confidence intervals leave little evidence that primary timing affects different types of primaries, or the two parties differently during the 1984-2016 period.

[Table 2 and Figure 14 about here]

Figure 14d (see also Table 2d) plots predicted turnout rates by primary month and party, over time. The panel on the left is restricted to primaries from the 1980s, and the panel to the right is restricted to primaries in the 2010s. Several interesting patterns emerge. First, the figure captures the overall decline in primary turnout parties have experienced over the past three decades; the highest instances of predicted turnout (by month) in the 2010s are still lower than the lowest instances of predicted turnout (by month) in the 1980s. Second, while predicted turnout rates in the two parties clearly track each other in both periods, in recent years, they have converged, and the curvilinear nature of the relationship between timing and turnout has become more prominent.

Finally, Figure 14e (multivariate results not shown) presents predicted turnout by party under two conditions: when the House primary is not held on the same day as the presidential primary (left panel) and when the House primary is held on the same day as the presidential primary (right panel). Consistent with the descriptive data presented above, predicted House primary turnout in both parties is higher in primaries that take place on the same day as the presidential primary. Additionally, same day primaries exhibit a moderate decline in turnout as the primary season progresses, again corroborating the patterns we observed in the bivariate. This is consistent with a process in which voters stop coming to the polls as some presidential aspirants begin to drop out of the race as others sew up their parties' nominations.

Cost and Competitiveness

The analyses of the cost and competitiveness of primaries follow the same basic pattern as the analysis of turnout that appears in the previous section. First, we present the direct relationship between timing and competitiveness (measured here using the fractionalization index described in the data section), and then we continue our search for patterns by breaking the analysis out by party and primary type and then by party and decade. The information presented in Tables 3 and 4 and Figures 15 and 16 leads us to the general conclusion that there is little evidence of a systematic relationship between primary timing and either cost or competitiveness.

[Table 3 and Figure 15 about here]

Turning first to cost, the various models perform well, with overall goodness-of-fit (R^2) ranging from .3 to .5. Figure 15a plots the predicted values of total spending over the course of the primary cycle, broken out by primary type and using the continuous "days since the first primary of the cycle" measure of timing.[17] We discern no pattern in these results, suggesting that primary spending varies independent of primary timing regardless of one's choice of how to operationalize timing. The predicted values in Figure 15b may dip slightly in the final month of primary season, but on the whole, patterns among all three types of primaries track each other and the range varies little over the course of the primary season. When we break the analysis out by primary type and party (see Figure 15c), we see that Republican spending in primaries appears to drop off toward the end of the cycle, while Democratic spending remains relatively steady. Finally, Figure 15 reveals little in the way of either intra-year effects of primary timing on spending or differences between the two parties. Indeed, Democratic and Republican primaries track each other when it comes to total primary spending.

Figure 16 (and Table 4) investigate competitiveness as a function of the various factors used in previous models. Incumbent primaries are, unsurprisingly, the least competitive of the three types, but again, neither the continuous nor the monthly operationalization of primary timing reveals a connection between timing and competitiveness. This is also true when we separate out incumbent primaries and model the likelihood that an incumbent will be in a "competitive" primary, operationalized here as one in which he or she receives less than 75 percent of the primary vote. Figure 16e demonstrates that the predicted probability of this happening hovers between 60 percent and 70 percent over the course of the primary season, with large confidence intervals. Breaking the analysis out by party and primary type over the entire 1984-2016 period again reveals two parties whose primary competitiveness varies little over the course of the primary season. Again, it is not until we break the analysis out by decade that we begin to notice variation in the way timing affects our outcome of interest (in this case competitiveness). In the 2010s, Democratic primaries appear to become less competitive as the primary season progresses, whereas Republican primaries appear to hold steady or even become slightly more competitive.

[Table 4 and Figure 16 about here]

[17] For space reasons, the results of the model from which these values and those that appear in Figure 16e are not presented here.

In conclusion, once we control for district, state, year, partisan effects, as well as other factors, our results reveal very little evidence for a linear relationship between primary timing and any of our dependent variables of interest. On the whole, the multivariate results corroborate the null findings presented above with regard to cost and competitiveness, and they call into question the modest effect of timing on turnout we identified in the bivariate analyses discussed in the previous section. This last point is particularly important since so many of the normative discussions of primary timing involve turnout. Our investigation reveals that conclusions about the relationship between primary timing and primary turnout are highly sensitive to model specification and analysts' choices about both appropriate baselines and how to operationalize primary timing. Simply including a variable capturing the number of days between the primary election and the general election in one's model, or interacting such a counter variable with primary type, can lead analysts to conclude that states wishing to maximize primary turnout might want to move their primaries as early as possible. This would be misleading. As we have demonstrated here, this downward trend vanishes once primary elections are aggregated by month and is instead replace, at least in recent years, by a curvilinear trend in which turnout dips in the summer before recovering in the fall.

The multivariate results here still do provide a slight suggestion that primary timing has had different effects on primaries held since 2010 than was the case in previous years. The rebound in turnout within both parties in late primaries and the differences between the two parties in the competitiveness of late primaries are worth investigating, but it is far too soon to conclude that something fundamental about the nature of primaries, or of the two parties' approach to primaries, has changed.

State Effects

Which States Changed their Primary Dates?

One important question in this analysis is whether any perceived timing effects are actually driven by characteristics of individual states. While the models presented in the previous sections are specified to account for such state-specific effects, as Figure 17 shows, there are noteworthy differences in primary timing according to region of the country. In each election year, at least three of the five New England states have held their primaries in the fall, while Plains states have favored summer primaries and the larger Midwestern industrial states have held primaries in the spring. Sawtooth patterns for some regions of the country are indicative of states that have tended to hold early concurrent presidential and congressional primaries but have held congressional primaries later in midterm election years.

[Figure 17 about here]

It is rare for states to change their primary dates. If one measures the mean number of days between the primary and general elections, twenty-one states have a standard deviation of over two weeks. Yet the standard deviations here generally capture three different patterns:

states such as Idaho or Connecticut, which moved their primary dates once during the 1978 to 2014 time period and then kept it at its new date; states such as Ohio, which hold their primaries earlier in presidential years than in nonpresidential years; and states that moved their primary once for idiosyncratic reasons and them moved it back, as Arkansas did in 1988. In a few cases, states did more than one of these things: Alabama, for instance, moved its primary from September to June in 1984, and has kept it there for every subsequent election except for one. Only seven states (Colorado, Georgia, Nevada, North Carolina, Oklahoma, Pennsylvania, and Washington) have changed their primary timing laws more than twice.

The general stability here stands in contrast to the volume of efforts to change primary dates; as we have discussed in greater detail elsewhere (Boatright 2016), 76 different bills to adjust state primary dates were introduced in 31 different states between 2000 and 2014. Six of these proposals were successful. We cannot infer the rationale for changes in primary dates from the data shown here. Yet our investigation of news coverage of some of these proposals yielded a variety of motives: increasing voter participation, improving voter knowledge of the candidates, and enhancing the ability of overseas military personnel to vote (all in support of establishing earlier primary dates); and decreasing interest group activity, decreasing campaign spending, allowing for the use of public schools as polling locations, and giving legislators more time to campaign in newly drawn districts (all in support of establishing later primary dates). These motives show a mixture of impressionistic claims about primary timing and state-specific concerns about unusual circumstances.

The overall stability in primary dates, especially among the states with fall primaries, shows that we must be careful to not to attribute any variation in competitiveness or turnout to primary timing effects as opposed to state-specific characteristics. In instances where states have moved their primaries, we can look at the relationship between fractionalization or turnout and a state's primary date. In California, the number of days between the primary and the general election is negatively correlated with fractionalization in primaries, while in Pennsylvania (a state which has changed its primary date frequently) turnout is positively correlated with the number of days between the primary and the general election. The change in primary date is not necessarily causal in either case.

Within-State Effects of Primary Date Change

One way to measure the effects of changes in primary date while controlling for state effects is to look at the experience of individual states. Our variables of interest here, turnout and fractionalization, may be affected from one year to the next within a state by a variety of contextual factors, including the presence or absence of a competitive statewide primary, differences in the number of open seats or incumbent candidates, and/or less measurable characteristics of the candidates who emerge.[18] Some of these factors may, in turn, be influenced in various ways by the primary date itself. Nonetheless, given that the stated goals of changes in primary date often have to do with influencing election outcomes or participation, it is worth looking at the trajectory of such things in states where the date changed.

[18] Because of the small numbers of cases in some of these states, we do not provide similar estimates for spending.

Many primary date changes are brought about by short-term problems; for instance, North Carolina separated its House and statewide primaries in 2016; it had moved its primary up from May 6 in 2014 to March 15 in 2016, but was forced to hold the House primary on June 7 as the result of a February court decision invalidating two of the districts. Other state have had similar one-time changes in primary date, caused by court rulings or other occurrences that were not a matter of the legislature's deliberate choice to change the primary date. Five states (Maryland, Mississippi, North Carolina, Ohio, and Pennsylvania) also hold their primaries in different months in presidential years than they do in nonpresidential years. Excluding the one-time changes and changes prompted by a desire to hold simultaneous presidential and state, there were eight states that made a lasting change of the date of the primary by more than one month at some point during the 1984-2016 time period. All of these changes involved moving the primary earlier in the cycle. Four of these states changed their primary date once. Connecticut held its primary in September in all but one year prior to 2004, but moved the date to early August in 2004. Hawaii moved its primary from mid-September to early August in 2012. Minnesota changed its primary date from September to August in 2010. And Utah moved its primary date from September to late June in 1994.

Four other states present slightly more complicated trajectories. California has usually held its primary in June, but it held the primary in March in 1996, 2000, 2002, and 2004. The move in 2000 was prompted by the state's change from a partisan primary to a nonpartisan blanket primary. The Supreme Court invalidated the blanket primary shortly before the 2002 primary, and the move in 2006 back to a June primary date was in part a consequence of the return to holding partisan primaries.[19] From 1978 to 1994, Ohio held its primary in either May or June, but beginning in 1996 it kept its midterm year primaries in May but held its statewide and presidential primaries in March during presidential election years. Colorado changed its primary date twice, from September to August in 1986 and from August to late June in 2012. And Nevada moved its primary date from September to August in 2006, and from August to June in 2010.

Table 5 shows a simple before-and-after comparison of fractionalization and turnout for all of these states, separating out the complicated cases of California and Ohio. We list the seasons in reverse order to make it easier for the reader to see the direction of change – in all of these cases, the change in date made the primary earlier, so one can see the results of the change by comparing numbers as one reads from left to right. The table shows that for the most part, primary fractionalization increases in these states following the move to an earlier primary, and turnout decreases. In nine of fourteen instances of primary date change across seasons (counting each party's primary as a case, and counting Nevada's primaries twice since the date was changed from fall to summer and then from summer to fall), fractionalization increases. Turnout decreases in thirteen of fourteen cases, although measures are complicated by variation among the states in whether primaries are held for uncontested nominations. The number of cases in some of these states is rather small, so we cannot be certain that we are not seeing idiosyncrasies brought about by the candidates or other attributes. There is also a secular trend toward lower primary participation, as noted earlier, so we cannot be certain that the change in date caused the decline, but the change certainly has not increased voter participation. We would

[19] For discussion of the California case, see the essays in Cain and Gerber 2002.

note, as well, that this is not a function of changing partisanship in these states. In most of these states, fractionalization increases within both parties following the date change.

[Table 5 about here]

Ohio and California are shown at the bottom of the table because of their different circumstances. The date change for these states does not technically fit our framework for defining seasons (all of the primaries held in these states were spring primaries according to our coding), but we have listed the May/early June primaries in these states as summer primaries for the purpose of fitting them into the same table as the other states. The Ohio primary does not seem remarkably different when it is held in March during the presidential years from 1996 to the present. Differences in both fractionalization and turnout are small and probably driven by factors other than the date change. In California, the four election cycles where the primary was held earlier exhibited substantially less competition than the years when primaries were held in June. This difference is not a consequence of the use of the blanket primary in 2000; even when that year is excluded, the three remaining years with March primaries still display similarly low levels of fractionalization.

In order to see the relationship between primary timing change and secular changes in fractionalization or competitiveness, it is useful to benchmark states against other states that might experience similar patterns. Using the ICPSR Region categories, we show some time series that illustrate the minimal effects of timing change. Figure 18 shows changes in turnout in Connecticut, compared to other New England states, and in Colorado, compared to other Mountain states. Connecticut moved its primary from fall to summer in 2004, and Colorado moved its primary from fall to summer in 2012. Neither state has a large enough congressional delegation that one can be certain changes in turnout are not simply the result of idiosyncratic factors. It is evident here, however, that the changes in primary timing have not caused these states to move in a different direction than their neighbors. Similar graphs for the other states discussed in Table 2 (not shown) also do not show effects, nor do graphs broken out by primary type (incumbent/challenger/open seat primary).

[Figure 18 about here]

One could, of course, single out other instances of changes in primary timing. The states here, however, provide the most dramatic before-and-after effects. In regards to turnout, the effects are not particularly noteworthy. Moving up the primary date did not increase turnout. Fractionalization did increase in many of these states, suggesting either that earlier primaries exhibit more competition or that fall primaries tend to be less competitive than primaries held earlier in the year.

It has been alleged in some discussions of primary law reform that changes in voter or candidate behavior take place a few elections after the change in primary rules; such claims have been made about the potential effects of California's recent reforms. It has also been argued that there can be notable one-election changes in response to changes in law. North Carolina's one-off change in primary timing in 2016, for instance, might have been expected to produce some change. Claims have even been made that some primary law changes have been enacted in order

to benefit one specific candidate. No lasting effects at the state level are evident, however, among the dependent variables we have considered here.

All of this should not be taken to diminish the importance and merit of other reasons for changes in primary dates. For instance, many states moved their primaries earlier in 2010 and subsequent years in order to comply with the Military and Overseas Voter Empowerment Act of 2009 and its new regulations regarding the distribution and counting of absentee ballots.[20] The proper counting of such ballot is arguably a legitimate goal in itself, apart from any effect it might have on voting behavior or election characteristics and outcomes. Other normative criteria have been presented at times regarding changes to primary dates; we should not assume all changes have to do with efforts to influence the elections themselves.

Discussion and Conclusions

In this paper, we have explored the relationship between the timing of congressional primaries and three dependent variables: turnout, cost, and competitiveness. A handful of extant studies have investigated some of these relationships (with mixed results), but our data allow us to extend the analysis across a larger range of elections and to control for potentially confounding variables others have omitted. Our results point us towards some potentially fruitful lines of inquiry, on the one hand, and help us identify areas that may prove less fruitful on the other.

Let us first address our findings on the cumulative 1984-2016 time series. Perhaps the most consistent message in our data is that despite occasional claims by reformers that the later primaries would reduce the cost of congressional elections, the timing of congressional primaries does not appear to affect the cost of either general elections or primary elections of any type. Nonetheless, we have collected campaign finance data on various types of contributions, as well as data on contributions to individual candidates from parties and interest groups. Even if the non-results on overall spending withstand more rigorous testing, there may be certain types of funds or candidates whose financial fortunes are affected by the date of the primary. There are many more dynamics to consider before we are prepared to close the book on the relationship between primary timing and the financing of primary and general elections.

In terms of competitiveness and turnout, the results are more mixed. Our seasonal analysis of variation in primary dates indicates that early primaries tend to be less competitive than late primaries, regardless of type. Yet in contrast, in nine of fourteen instances we investigated, moving the primary earlier in the year resulted in an uptick in competitiveness *within* a state. Our multivariate analysis, however, revealed no substantively meaningful connection between primary timing and competitiveness (see Figure 16b). These discrepant results merit further exploration.

[20] For discussion see the Pew Foundation issue brief on the MOVE Act, at http://www.pewtrusts.org/~/media/legacy/uploadedfiles/pcs_assets/2010/militaryoverseasvotingbrief1pdf.pdf.

We also uncovered evidence that speaks to reformers' claims about the relationship between voter participation and primary timing. Our seasonal data reveal that participation in late primaries tends to be slightly lower, on average, than participation in earlier primaries, although with the exception of incumbent primaries, the substantive differences are negligible and appear to almost certainly be driven by concurrence with the presidential primary. However, our multivariate analysis of turnout reveals that the relationship between timing and turnout may be changing. In recent years, turnout has dipped in the summer before recovering in the fall. Finally, our brief analysis of general election turnout suggests that later primaries tend to be associated with slightly *higher* turnout on average. This may be driven by the characteristics of the specific states that hold these late primaries, but it is also possible that later primaries and the attention they draw boost the effectiveness of mobilization efforts of candidates, campaigns, and groups in the general election. We still have some work to do in order to fully understand the interplay or spillover effects of primary and general election mobilization efforts.

Despite the lack of aggregate effects, there is some evidence that the individual states that have moved their primaries have experienced lower turnout but greater competition. It is hard to know whether these changes have resulted from idiosyncrasies within these states, or whether they have merely been state-level manifestations of a natural trend toward higher fractionalization and lower turnout. Yet the persistence of the notion that primary date changes will bring about results continues to drive legislative proposals within the states. We are hopeful that this research will serve to inform such efforts.

We are also struck by the difference between our comparisons of primaries during the 2010 to 2016 period, on the one hand, and primaries during the 1980s. Our interest in undertaking this project has been in testing a set of claims that were made a century or more ago, about the innate structural characteristics of primary elections. It is far too soon to contend that anything has changed in the conduct of primaries, and our evidence for making such claims is slight. Nonetheless, some differences are noteworthy.

These differences speak to two major concerns in contemporary literature on congressional elections. They speak, first of all, to emerging differences between the two parties. To recap, in our full time series, turnout appeared to hold steady or decline in later primaries. But during the 2010s, a curvilinear pattern seems to have emerged in which turnout in both parties' primaries dips in the summer before recovering in the fall. Democratic candidates' spending holds steady throughout the cycle while Democratic primaries become slightly less competitive. Similarly, Republican spending and competition remain about the same in late primaries. All of these factors, when put together, correspond to what Boatright (2013) has described as the "nationalization" of primaries, and they correspond as well to research on differences between the parties (e.g. Grossmann and Hopkins 2016). Democrats have become a less fractured party than they were believed to be during prior decades, and they have become more adept at ironing out internal differences in order to prepare for the general election. Republicans, on the other hand, have exhibited more difficulty in settling upon "party preferred" candidates, and the consistency in turnout and lack of obvious preparation for the general election in the late primaries is indicative of this. The 2010-2016 time period was, on balance, a period of slight Republican advantage – Democrats picked up seats in two of these election years (2012 and 2016) while Republicans gained seats in 2010 and 2014. Republicans won more seats

overall during this period, however, suggesting that the differences between the parties in how primaries played out did not obviously harm Republicans.

Second, these differences have relevance to contemporary analyses of political polarization. Many studies of primary elections have been prompted by the increasing ideological gap between the parties and the perception that primary elections have pulled nominees away from the political center. Does the primary calendar play a role in this? Or, to reframe the question, would modifications to the primary calendar lead to less extreme nominees? These are, unfortunately, questions we cannot directly answer using the data we have here. Logic suggests that more fractionalized primaries can produce more extreme nominees – a plurality winner is more likely to be at one extreme or the other than a candidate who wins a majority of the votes. Yet most studies of polarization have expressed more concern about extremism on the part of Republicans than on the part of Democrats (e.g. Mann and Ornstein 2012), and there is no evidence that fractionalization among Republicans is lower at any point in the calendar. The causal links between turnout and spending and polarization are far more tenuous, and a case can be made that any number of different patterns might influence polarization.

We included in the data we gathered a variety of measures of candidate ideology, including NOMINATE and CF scores. However, the real polarization-related question here has to do with our ability to distinguish between "insider" and "outsider" candidates, or between party-preferred candidates and insurgents. Such differences are obvious in incumbent primaries, but incumbents are opposed rarely enough that it is not clear that large N analyses are the most fruitful way to advance our understanding of the effect of primary timing on incumbents. For every story about an incumbent who faces a serious challenge from a more ideologically extreme candidate in an early primary (states such as Texas or Illinois) there is another about a surprising challenge in a September primary. It is hard to distinguish between timing and features of the states in which these primaries were held or characteristics of the individual candidates. Nonetheless, we uncovered no evidence that the probability of an incumbent receiving less than 75% of the primary vote varies over the course of the primary season. In other types of races, we can measure ideology but it is still difficult to determine party preferences or to draw conclusions about the many factors that may have led to a less centrist candidate's victory. Nonetheless, the evidence here that timing is a major area of concern in this regard is scant.

We undertook this study in part to fill a lacuna in the literature on primaries. Of the many claims made in the 1910s and 1920s about primaries, the arguments about when primaries should happen have been, in our view, the claims that have received the least scrutiny. This may be in part because the effects of timing are so difficult to measure. Our evidence here suggests, however, that primary timing is not responsible for very much of the variation we seen in primary turnout, competitiveness, or spending. There are some signs that timing has had different effects on Democrats and Republicans over the past few years, and we should continue to monitor such developments. In sum, however, there may still be good normative reasons why states might wish to have their primaries at particular times, but we should be careful not to assume that the timing of primaries is a cause of any of the characteristics, good or bad, of our contemporary Congress.

Table 1: State Primary Dates, 2014 and 2016

State	2014	2016	State	2014	2016
Texas	March 1	**March 1**	Virginia	June 10	June 14
Illinois	March 18	**March 15**	Colorado	June 24	June 28
Indiana	May 6	**May 3**	Maryland	June 24	**April 26**
North Carolina	May 6	June 7*	New York	June 24	June 28
Ohio	May 6	**March 15**	Oklahoma	June 24	June 28
Nebraska	May 13	**May 10**	Utah	June 24	June 28
West Virginia	May 13	**May 10**	Kansas	August 5	August 2
Arkansas	May 20	**March 1**	Michigan	August 5	August 2
Georgia	May 20	May 24	Missouri	August 5	August 2
Idaho	May 20	May 17	Washington	August 5	August 2
Kentucky	May 20	**May 17**	Tennessee	August 7	August 4
Oregon	May 20	**May 17**	Hawaii	August 9	August 13
Pennsylvania	May 20	**April 26**	Connecticut	August 12	August 9
Alabama	June 3	**March 1**	Minnesota	August 12	August 9
California	June 3	**June 7**	Wisconsin	August 12	August 9
Iowa	June 3	June 7	Alaska	August 19	August 16
Mississippi	June 3	**March 8**	Wyoming	August 19	August 16
Montana	June 3	**June 7**	Arizona	August 26	August 30
New Jersey	June 3	**June 7**	Florida	August 26	August 30
New Mexico	June 3	**June 7**	Vermont	August 26	August 9
South Dakota	June 3	**June 7**	Delaware	September 9	September 13
Maine	June 10	June 14	Massachusetts	September 9	September 8
Nevada	June 10	June 14	New Hampshire	September 9	September 13
North Dakota	June 10	June 14	Rhode Island	September 9	September 13
South Carolina	June 10	June 14	Louisiana**	November 4	November 8

Note: States listed in order of their 2014 primary dates. Concurrent primaries listed in **bold.**
* Because of complications caused by a court-ordered redistricting, North Carolina held its congressional primaries on June 7, 2016, but its Senate primary concurrently with its presidential primary on March 15.
** Louisiana holds a nonpartisan "jungle" primary on the date of the general election; a runoff is held in December if no candidate receives 50 percent of the vote.

Table 2a: Effect of Primary Timing on Primary Turnout, 1984-2016

Independent Variables	Coefficients	Standard Errors
Primary Timing (Days Since First Primary)	-0.000130***	(0.0000314)
Challenger Primary	-0.0362***	(0.00356)
Open Seat Primary	-0.0133**	(0.00423)
Challenger Primary * Primary Timing	0.0000758***	(0.0000226)
Open Seat Primary * Primary Timing	0.0000432	(0.0000253)
State and Year Dummy Variables Omitted to Conserve Space		
Strong Traditional Party Org (Mayhew)	-0.0521***	(0.0151)
Total Spending (ln)	0.00694***	(0.000331)
Same Day Presidential Primary	0.0225***	(0.00241)
Same Day Gubernatorial Primary	0.00968***	(0.00145)
Republican Primary	-0.0134***	(0.00217)
Lagged Primary Turnout (%)	0.0761***	(0.0131)
Closed Primary	0.0188***	(0.00555)
Semi-Closed Primary	0.0219***	(0.00661)
Semi-Open Primary	0.0138*	(0.00641)
Constant	0.0415**	(0.0149)
N	5,178	
Overall R^2	.47	
Wald Chi Square$_{(df=70)}$	4,345.3***	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is turnout (%) in U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) and dummy variables for each year (1984 is the baseline) to account for effects unique to any given state or any given election cycle. Full results available upon request.

Table 2b: Effect of Primary Timing on Primary Turnout, by Primary Type, 1984-2016

Independent Variables	Coefficients	Standard Errors
*Primary Type * Primary Month Interactions*		
Inc * Mar	-0.0306**	(0.0115)
Inc * Apr	-0.0432***	(0.0123)
Inc * May	-0.0587***	(0.0117)
Inc * Jun	-0.0591***	(0.0114)
Inc * Jul	-0.0503***	(0.0131)
Inc * Aug	-0.0423***	(0.0118)
Inc * Sep	-0.0503***	(0.0113)
Chal * Feb	-0.0333**	(0.0123)
Chal * Mar	-0.0699***	(0.0114)
Chal * Apr	-0.0669***	(0.0115)
Chal * May	-0.0845***	(0.0120)
Chal * Jun	-0.0816***	(0.0116)
Chal * Jul	-0.100***	(0.0142)
Chal * Aug	-0.0714***	(0.0116)
Chal * Sep	-0.0671***	(0.0115)
Open * Feb	-0.0348*	(0.0160)
Open * Mar	-0.0453***	(0.0122)
Open * Apr	-0.0326*	(0.0128)
Open * May	-0.0658***	(0.0120)
Open * Jun	-0.0675***	(0.0118)
Open * Jul	-0.0511***	(0.0132)
Open * Aug	-0.0541***	(0.0118)
Open * Sep	-0.0535***	(0.0115)
State and Year Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	-0.0596***	(0.0153)
Total Spending (ln)	0.00691***	(0.000327)
Same Day Presidential Primary	0.0211***	(0.00242)
Same Day Gubernatorial Primary	0.0110***	(0.00150)
Republican Primary	-0.0135***	(0.00218)

Lagged Primary Turnout (%)	0.0692***	(0.0132)
Closed Primary	0.0158**	(0.00561)
Semi-Closed Primary	0.0195**	(0.00662)
Semi-Open Primary	0.0107	(0.00641)
Constant	0.0899***	(0.0178)
N	5,178	
Overall R^2	.48	
Wald Chi Square$_{(df=88)}$	4,931.6***	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is turnout (%) in U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. The baseline for the primary type * month interactions is an incumbent primary held in February. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) and dummy variables for each year (1984 is the baseline) to account for effects unique to any given state or any given election cycle. Full results available upon request.

Table 2c: Effect of Primary Timing on Primary Turnout, by Primary Type and Party, 1984-2016

Independent Variables	Coefficients	Standard Errors
Direct Effects Primary Party		
Republican Primary Incumbent Primary	-0.0415**	(0.0159)
Direct Effects Primary Type		
Challenger Primary	-0.0236	(0.0180)
Open Seat Primary	-0.0736***	(0.0140)
*Party * Primary Type Interactions*		
GOP Primary * Chal Primary	0.000151	(0.0201)
GOP Primary * Open Primary	0.0725***	(0.0170)
Direct Effects – Primary Month		
Mar	-0.0390**	(0.0146)
Apr	-0.0433**	(0.0156)
May	-0.0614***	(0.0144)
Jun	-0.0701***	(0.0138)
Jul	-0.0631***	(0.0158)
Aug	-0.0524***	(0.0148)
Sep	-0.0550***	(0.0134)
*Party * Month Interactions*		
GOP * Mar	0.0335*	(0.0166)
GOP * Apr	0.0160	(0.0192)
GOP * May	0.0226	(0.0165)
GOP * Jun	0.0396*	(0.0160)
GOP * Jul	0.0436*	(0.0195)
GOP * Aug	0.0356*	(0.0170)
GOP * Sep	0.0259	(0.0161)
*Primary Type * Primary Month Interactions*		
Chal * Mar	-0.00793	(0.0195)

Chal * Apr	-0.00874	(0.0207)
Chal * May	-0.00333	(0.0185)
Chal * Jun	0.00744	(0.0182)
Chal * Jul	-0.0314	(0.0226)
Chal * Aug	-0.00453	(0.0185)
Chal * Sept	0.0100	(0.0181)
Open * Mar	0.0626***	(0.0159)
Open * Apr	0.0680***	(0.0191)
Open * May	0.0649***	(0.0153)
Open * Jun	0.0677***	(0.0145)
Open * Jul	0.0884***	(0.0219)
Open * Aug	0.0591***	(0.0151)
Open * Sep	0.0792***	(0.0147)
*Primary Party * Type * Month Interactions*		
GOP * Chal * Mar	-0.0154	(0.0218)
GOP * Chal * Apr	0.0167	(0.0238)
GOP * Chal * May	0.00142	(0.0210)
GOP * Chal * Jun	-0.0138	(0.0206)
GOP * Chal * Jul	0.00245	(0.0270)
GOP * Chal * Aug	-0.00238	(0.0213)
GOP * Chal * Sep	-0.00646	(0.0206)
GOP * Open * Mar	-0.0808***	(0.0204)
GOP * Open * Apr	-0.0386	(0.0233)
GOP * Open * May	-0.0697***	(0.0184)
GOP * Open * Jun	-0.0777***	(0.0179)
GOP * Open * Jul	-0.105**	(0.0326)
GOP * Open * Aug	-0.0679***	(0.0188)
GOP * Open * Sep	-0.0918***	(0.0187)
State and Year Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	-0.0597***	(0.0156)
Total Spending (ln)	0.00695***	(0.000326)
Same Day Presidential Primary	0.0209***	(0.00248)
Same Day Gubernatorial Primary	0.0109***	(0.00150)
Lagged Primary Turnout (%)	0.0698***	(0.0134)
Closed Primary	0.0154**	(0.00565)
Semi-Closed Primary	0.0198**	(0.00662)
Semi-Open Primary	0.0116	(0.00645)

Constant	0.0955***	(0.0193)
N	5,178	
Overall R^2	.49	
Wald Chi Square$_{(df=111)}$	13,147.1***	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is turnout (%) in U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) and dummy variables for each year (1984 is the baseline) to account for effects unique to any given state or any given election cycle. Full results available upon request.

Table 2d: Effect of Primary Timing on Primary Turnout, by Primary Type and Party, 2010s v. pre-2010 period

Independent Variables	Coefficients	Standard Errors
Direct Effects (Party, Decade, Primary Month)		
GOP Primary	-0.0620***	(0.0101)
Decade 2010s	-0.0975***	(0.0112)
Mar	-0.0779***	(0.00824)
Apr	-0.0642***	(0.0123)
May	-0.0865***	(0.00945)
Jun	-0.0892***	(0.00922)
Jul	-0.102***	(0.0120)
Aug	-0.0863***	(0.0106)
Sep	-0.0782***	(0.00954)
*Party * Decade Interaction*		
GOP * 2010s	0.0871***	(0.0139)
*Party * Month Interactions*		
GOP * Mar	0.0350***	(0.00931)
GOP * Apr	0.0335*	(0.0148)
GOP * May	0.0299**	(0.0113)
GOP * Jun	0.0454***	(0.0103)
GOP * Jul	0.0295	(0.0154)
GOP * Aug	0.0397**	(0.0123)
GOP * Sep	0.0369***	(0.0105)
*Decade * Month Interactions*		
2010s * Mar	0.104***	(0.0139)
2010s * Apr	0.0986***	(0.0171)
2010s * May	0.0920***	(0.0121)
2010s * Jun	0.105***	(0.0120)
2010s * Jul	0.0976***	(0.0159)
2010s * Aug	0.108***	(0.0121)
2010s * Sep	0.103***	(0.0119)
*Party * Decade * Month Interactions*		

GOP * 2010s * Mar	-0.0323	(0.0172)
GOP * 2010s * Apr	-0.0631**	(0.0200)
GOP * 2010s * May	-0.0494**	(0.0151)
GOP * 2010s * Jun	-0.0718***	(0.0148)
GOP * 2010s * Jul	-0.0302	(0.0216)
GOP * 2010s * Aug	-0.0522***	(0.0155)
GOP * 2010s * Sept	-0.0559***	(0.0154)
State Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	-0.0568***	(0.0155)
Total Spending (ln)	0.00854***	(0.000347)
Same Day Presidential Primary	0.0155***	(0.00150)
Same Day Gubernatorial Primary	0.00926***	(0.00152)
Lagged Primary Turnout (%)	0.129***	(0.0118)
Election Year Counter	-0.00196***	(0.0000948)
Closed Primary	-0.00751	(0.00440)
Semi-Closed Primary	-0.00666	(0.00488)
Semi-Open Primary	-0.0128*	(0.00519)
Constant	4.003***	(0.191)
N	5,178	
Overall R^2	.49	
Wald Chi Square$_{(df=80)}$	4,630.6***	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is turnout (%) in U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) to account for effects unique to any given state. Full results available upon request.

Table 3a: Effect of Primary Timing on Cost of Primary, by Primary Type, 1984-2016

Independent Variables	Coefficients	Standard Errors
*Primary Type * Primary Month Interactions*		
Inc * Mar	-0.0680	(0.231)
Inc * Apr	0.215	(0.322)
Inc * May	-0.126	(0.261)
Inc * Jun	-0.0776	(0.250)
Inc * Jul	0.0686	(0.305)
Inc * Aug	0.0405	(0.270)
Inc * Sep	-0.139	(0.254)
Chal * Feb	-1.164*	(0.535)
Chal * Mar	-1.538***	(0.278)
Chal * Apr	-1.100**	(0.421)
Chal * May	-1.066***	(0.293)
Chal * Jun	-1.433***	(0.273)
Chal * Jul	-1.156**	(0.413)
Chal * Aug	-1.111***	(0.291)
Chal * Sep	-1.238***	(0.283)
Open * Feb	0.330	(0.494)
Open * Mar	0.452	(0.268)
Open * Apr	-0.204	(0.405)
Open * May	-0.137	(0.298)
Open * Jun	0.317	(0.277)
Open * Jul	0.146	(0.393)
Open * Aug	0.385	(0.288)
Open * Sep	0.331	(0.271)
State and Year Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	0.172	(0.525)
Lagged Total Spending (ln)	0.00694	(0.00558)
Same Day Presidential Primary	-0.0305	(0.0957)
Same Day Gubernatorial Primary	-0.0491	(0.0591)
Republican Primary	0.107	(0.0631)
Lagged Primary Turnout (%)	-0.283	(0.184)
Closed Primary	-0.282	(0.211)

Semi-Closed Primary	-0.259	(0.209)
Constant	11.92***	(0.403)
N	5,185	
Overall R^2	.29	
Wald Chi Square$_{(df=87)}$	2,069.3***	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is the natural log of total spending (all candidates) in U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. The baseline for the primary type * month interactions is an incumbent primary held in February. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) and dummy variables for each year (1984 is the baseline) to account for effects unique to any given state or any given election cycle. Full results available upon request.

Table 3b: Effect of Primary Timing on Cost of Primary, by Primary Type and Party, 1984-2016

Independent Variables	Coefficients	Standard Errors
Direct Effects Primary Party		
Republican Primary	0.388	(0.754)
Direct Effects Primary Type		
Challenger Primary	-1.042	(0.723)
Open Seat Primary	0.371	(0.241)
*Party * Primary Type Interactions*		
GOP Primary * Chal Primary	-0.407	(1.406)
GOP Primary * Open Primary	-0.258	(0.983)
Direct Effects – Primary Month		
Mar	-0.0266	(0.223)
Apr	0.344	(0.310)
May	-0.0737	(0.247)
Jun	-0.0152	(0.242)
Jul	0.122	(0.318)
Aug	0.203	(0.267)
Sep	0.0382	(0.236)
*Party * Month Interactions*		
GOP * Mar	-0.218	(0.777)
GOP * Apr	-0.448	(0.807)
GOP * May	-0.246	(0.761)
GOP * Jun	-0.254	(0.758)
GOP * Jul	-0.242	(0.786)
GOP * Aug	-0.502	(0.765)
GOP * Sep	-0.621	(0.758)
*Primary Type * Primary Month Interactions*		
Chal * Mar	-0.604	(0.709)
Chal * Apr	0.0417	(0.847)
Chal * May	0.153	(0.745)

Chal * Jun	-0.321	(0.732)
Chal * Jul	-0.260	(0.892)
Chal * Aug	-0.316	(0.743)
Chal * Sept	-0.120	(0.752)
Open * Mar	0.304	(0.311)
Open * Apr	-0.355	(0.510)
Open * May	-0.517	(0.324)
Open * Jun	-0.0703	(0.285)
Open * Jul	-0.215	(0.490)
Open * Aug	-0.242	(0.326)
Open * Sep	0.0486	(0.300)
*Primary Party * Type * Month Interactions*		
GOP * Chal * Mar	0.694	(1.428)
GOP * Chal * Apr	-0.153	(1.530)
GOP * Chal * May	0.285	(1.423)
GOP * Chal * Jun	0.409	(1.415)
GOP * Chal * Jul	0.502	(1.519)
GOP * Chal * Aug	0.789	(1.430)
GOP * Chal * Sep	0.637	(1.456)
GOP * Open * Mar	-0.0443	(0.996)
GOP * Open * Apr	-0.651	(1.361)
GOP * Open * May	0.514	(1.037)
GOP * Open * Jun	0.446	(1.007)
GOP * Open * Jul	0.0884	(1.212)
GOP * Open * Aug	0.655	(1.031)
GOP * Open * Sep	0.449	(1.011)
State and Year Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	0.151	(0.522)
Lagged Total Spending (ln)	0.00661	(0.00562)
Same Day Presidential Primary	-0.0281	(0.0962)
Same Day Gubernatorial Primary	-0.0502	(0.0592)
Lagged Primary Turnout (%)	-0.271	(0.184)
Closed Primary	-0.260	(0.211)
Semi-Closed Primary	-0.245	(0.211)
Semi-Open Primary	0.151	(0.522)
Constant	11.85***	(0.383)

N	5,185
Overall R^2	.30
Wald Chi Square$_{(df=110)}$	52,425.6***

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is the natural log of total spending (all candidates) in U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) and dummy variables for each year (1984 is the baseline) to account for effects unique to any given state or any given election cycle. Full results available upon request.

Table 3c: Effect of Primary Timing on Cost of Primary, by Primary Type and Party, 2010s v. pre-2010 period

Independent Variables	Coefficients	Standard Errors
Direct Effects (Party, Decade, Primary Month)		
GOP Primary	0.0215	(0.646)
Decade 2010s	-0.702	(0.570)
Mar	-0.462	(0.261)
Apr	-0.164	(0.361)
May	-0.189	(0.292)
Jun	-0.452	(0.279)
Jul	0.0873	(0.365)
Aug	0.0170	(0.314)
Sep	0.000499	(0.269)
*Party * Decade Interaction*		
GOP * 2010s	-0.108	(0.864)
*Party * Month Interactions*		
GOP * Mar	0.0612	(0.673)
GOP * Apr	-0.642	(0.781)
GOP * May	-0.207	(0.649)
GOP * Jun	-0.0183	(0.660)
GOP * Jul	-0.129	(0.705)
GOP * Aug	-0.130	(0.680)
GOP * Sep	-0.104	(0.667)
*Decade * Month Interactions*		
2010s * Mar	-0.233	(0.534)
2010s * Apr	0.833	(0.647)
2010s * May	-0.0705	(0.606)
2010s * Jun	0.377	(0.599)
2010s * Jul	-0.139	(0.716)
2010s * Aug	0.116	(0.620)
2010s * Sep	0.873	(0.591)
*Party * Decade * Month Interactions*		
GOP * 2010s * Mar	0.582	(0.928)

GOP * 2010s * Apr	-0.117	(1.244)
GOP * 2010s * May	0.808	(0.900)
GOP * 2010s * Jun	0.140	(0.894)
GOP * 2010s * Jul	0.429	(0.992)
GOP * 2010s * Aug	0.441	(0.908)
GOP * 2010s * Sept	-0.595	(0.893)
State Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	0.129	(0.537)
Lagged Total Spending (ln)	0.0540***	(0.00470)
Same Day Presidential Primary	0.0101	(0.0557)
Same Day Gubernatorial Primary	0.0680	(0.0568)
Competitiveness (Fractionalization)	1.683***	(0.155)
Election Year Counter	0.0613***	(0.00427)
Closed Primary	-0.0642	(0.189)
Semi-Closed Primary	-0.0781	(0.220)
Semi-Open Primary	-0.0555	(0.218)
Constant	-111.0***	(8.632)
N	5,185	
Overall R^2	.49	
Wald Chi Square$_{(df=80)}$	2,110.4***	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is the natural log of total spending (all candidates) in U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) to account for effects unique to any given state. Full results available upon request.

Table 4a: Effect of Primary Timing on Competitiveness of Primary, by Primary Type, 1984-2016

Independent Variables	Coefficients	Standard Errors
*Primary Type * Primary Month Interactions*		
Inc * Mar	0.0378	(0.0293)
Inc * Apr	0.0900*	(0.0394)
Inc * May	0.0385	(0.0316)
Inc * Jun	0.0572	(0.0323)
Inc * Jul	0.0795*	(0.0386)
Inc * Aug	0.0251	(0.0346)
Inc * Sep	0.0601	(0.0339)
Chal * Feb	0.302***	(0.0537)
Chal * Mar	0.311***	(0.0290)
Chal * Apr	0.266***	(0.0403)
Chal * May	0.299***	(0.0324)
Chal * Jun	0.297***	(0.0323)
Chal * Jul	0.264***	(0.0380)
Chal * Aug	0.293***	(0.0348)
Chal * Sep	0.288***	(0.0350)
Open * Feb	0.274***	(0.0493)
Open * Mar	0.321***	(0.0312)
Open * Apr	0.404***	(0.0458)
Open * May	0.344***	(0.0335)
Open * Jun	0.354***	(0.0332)
Open * Jul	0.297***	(0.0393)
Open * Aug	0.312***	(0.0356)
Open * Sep	0.310***	(0.0357)
State and Year Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	0.0106	(0.0261)
Total Spending (ln)	0.0133***	(0.00171)
Turnout	0.761***	(0.0612)
Republican Primary	0.0299***	(0.00560)
Lagged Fractionalization	0.0410***	(0.00977)

Closed Primary	0.0492*	(0.0219)
Semi-Closed Primary	0.0605*	(0.0245)
Semi-Open Primary	0.0481*	(0.0241)
Constant	-0.0754	(0.0525)
N	5,178	
Overall R^2	.50	
Wald Chi Square$_{(df=87)}$	5,955.80***	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is an index of fractionalization in which higher values indicate more competitive elections (see text for details) in all contested U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. Baseline for primary type * primary month interactions is an incumbent primary held in February. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) and dummy variables for each year (1984 is the baseline) to account for effects unique to any given state or any given election cycle. Full results available upon request.

Table 4b: Effect of Primary Timing on Competitiveness of Primary, by Primary Type and Party, 1984-2016

Independent Variables	Coefficients	Standard Errors
Direct Effects Primary Party		
Republican Primary	0.150*	(0.0736)
Direct Effects Primary Type		
Challenger Primary	0.274***	(0.0712)
Open Seat Primary	0.270***	(0.0331)
*Party * Primary Type Interactions*		
GOP Primary * Chal Primary	-0.0305	(0.110)
GOP Primary * Open Primary	-0.0851	(0.0956)
Direct Effects – Primary Month		
Mar	0.0707*	(0.0335)
Apr	0.114**	(0.0428)
May	0.0532	(0.0346)
Jun	0.0644	(0.0356)
Jul	0.126**	(0.0441)
Aug	0.0473	(0.0384)
Sep	0.0830*	(0.0364)
*Party * Month Interactions*		
GOP * Mar	-0.141*	(0.0703)
GOP * Apr	-0.123	(0.0915)
GOP * May	-0.104	(0.0749)
GOP * Jun	-0.0882	(0.0755)
GOP * Jul	-0.176*	(0.0807)
GOP * Aug	-0.119	(0.0757)
GOP * Sep	-0.129	(0.0736)
*Primary Type * Primary Month Interactions*		
Chal * Mar	-0.00686	(0.0685)
Chal * Apr	-0.107	(0.0763)
Chal * May	-0.00565	(0.0727)

Chal * Jun	-0.0138	(0.0730)
Chal * Jul	-0.114	(0.0774)
Chal * Aug	-0.00820	(0.0738)
Chal * Sept	-0.0184	(0.0741)
Open * Mar	-0.000227	(0.0427)
Open * Apr	0.0238	(0.0486)
Open * May	0.0524	(0.0369)
Open * Jun	0.0197	(0.0383)
Open * Jul	-0.109	(0.0567)
Open * Aug	0.0148	(0.0384)
Open * Sep	-0.0149	(0.0374)
*Primary Party * Type * Month Interactions*		
GOP * Chal * Mar	0.0435	(0.107)
GOP * Chal * Apr	0.0490	(0.130)
GOP * Chal * May	0.0130	(0.113)
GOP * Chal * Jun	-0.0116	(0.113)
GOP * Chal * Jul	0.0905	(0.120)
GOP * Chal * Aug	0.0332	(0.113)
GOP * Chal * Sep	-0.00824	(0.113)
GOP * Open * Mar	0.109	(0.0939)
GOP * Open * Apr	0.129	(0.119)
GOP * Open * May	0.0507	(0.0982)
GOP * Open * Jun	0.0974	(0.0991)
GOP * Open * Jul	0.202	(0.120)
GOP * Open * Aug	0.0888	(0.0987)
GOP * Open * Sep	0.0797	(0.0985)
State and Year Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	0.00724	(0.0264)
Total Spending (ln)	0.0133***	(0.00171)
Lagged Fractionalization	0.0405***	(0.00975)
Turnout (%)	0.752***	(0.0621)
Closed Primary	0.0484*	(0.0220)
Semi-Closed Primary	0.0597*	(0.0246)
Semi-Open Primary	0.0483*	(0.0244)
Constant	-0.0954	(0.0539)

N	5,178
Overall R^2	.50
Wald Chi Square$_{(df=110)}$	10,704.4***

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is an index of fractionalization in which higher values indicate more competitive elections (see text for details) in all contested U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) and dummy variables for each year (1984 is the baseline) to account for effects unique to any given state or any given election cycle. Full results available upon request.

Table 4c: Effect of Primary Timing on Competitiveness of Primary, by Primary Type and Party, 2010s v. pre-2010 period

Independent Variables	Coefficients	Standard Errors
Direct Effects (Party, Decade, Primary Month)		
GOP Primary	0.111	(0.0665)
Decade 2010s	-0.00775	(0.1000)
Mar	0.0307	(0.0424)
Apr	0.174***	(0.0504)
May	-0.0116	(0.0453)
Jun	0.0228	(0.0454)
Jul	0.00662	(0.0501)
Aug	-0.0232	(0.0489)
Sep	0.0238	(0.0460)
*Party * Decade Interaction*		
GOP * 2010s	0.0585	(0.118)
*Party * Month Interactions*		
GOP * Mar	-0.0950	(0.0658)
GOP * Apr	-0.0273	(0.0714)
GOP * May	-0.0820	(0.0671)
GOP * Jun	-0.0595	(0.0679)
GOP * Jul	-0.0880	(0.0738)
GOP * Aug	-0.0486	(0.0705)
GOP * Sep	-0.0925	(0.0691)
*Decade * Month Interactions*		
2010s * Mar	0.0407	(0.104)
2010s * Apr	-0.159	(0.109)
2010s * May	0.0726	(0.102)
2010s * Jun	0.0355	(0.102)
2010s * Jul	-0.0180	(0.102)
2010s * Aug	0.0830	(0.103)
2010s * Sep	-0.0289	(0.105)
*Party * Decade * Month Interactions*		

GOP * 2010s * Mar	-0.0667	(0.133)
GOP * 2010s * Apr	-0.109	(0.134)
GOP * 2010s * May	-0.0565	(0.121)
GOP * 2010s * Jun	-0.0439	(0.122)
GOP * 2010s * Jul	0.0708	(0.123)
GOP * 2010s * Aug	-0.0728	(0.122)
GOP * 2010s * Sept	0.0997	(0.130)
State Dummy Variables Omitted to Conserve Space		
Controls		
Strong Traditional Party Org (Mayhew)	0.0528	(0.0315)
Lagged Competitiveness (Fractionalization)	-0.0718***	(0.0120)
Turnout (%)	0.0788	(0.0719)
Total Spending (ln)	0.00833***	(0.00201)
Election Year Counter	-0.00241***	(0.000567)
Closed Primary	0.0104	(0.0234)
Semi-Closed Primary	0.0671*	(0.0291)
Semi-Open Primary	0.0497	(0.0277)
Constant	5.100***	(1.138)
N	5,178	
Overall R^2	.19	
Wald Chi Square$_{(df=79)}$	1,506.6***	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is an index of fractionalization in which higher values indicate more competitive elections (see text for details) in all contested U.S. House primary elections, 1984-2016. Numbers in second column are random effects GLS regression coefficients. Robust standard errors, clustered by a variable that uniquely identifies the party holding the primary in a given district, are reported in the third column. Model includes dummy variables for all multi-district states (the seven single district states serve as the baseline) to account for effects unique to any given state. Full results available upon request.

Table 2: Effects of Primary Date Changes, Selected States

	Year	***Fractionalization***			***Turnout (%)***		
		Fall	*Summer*	*Spring*	*Fall*	*Summer*	*Spring*
Colorado	1986						
Democratic		.043 (22)	.066 (95)		6.26 (21)	5.04 (88)	
Republican		.154 (22)	.199 (95)		7.95 (22)	7.25 (91)	
Connecticut	2004						
Democratic		.095 (54)	.056 (24)		5.82 (10)	3.86 (3)	
Republican		.088 (52)	.153 (26)		4.01 (17)	4.06 (7)	
Hawaii	2012						
Democratic		.245 (33)	.357 (6)		22.96 (26)	19.88 (4)	
Republican		.319 (29)	.325 (6)		5.37 (23)	3.64 (4)	
Minnesota	2010						
Democratic		.139 (128)	.186 (30)		9.90 (108)	6. 62 (16)	
Republican		.125 (126)	.198 (30)		5.75 (96)	3.43 (16)	
Nevada	2006, 2010						
Democratic		.455 (7)	.266 (22)	.399 (15)	14.58 (7)	9.43 (14)	3.87 (8)
Republican		.466 (7)	.360 (25)	.464 (15)	10.84 (5)	9.35 (19)	5.24 (8)
Utah	1994						
Democratic		.187 (8)	.072 (36)		8.20 (2)	3.46 (4)	
Republican		.188 (10)	.167 (39)		16.16 (4)	10.77 (13)	
California	1996, 2000-2004						
Democratic			.200 (772)	.118 (210)		12.11 (719)	11.10 (207)
Republican			.231 (771)	.161 (210)		9.88 (694)	10.29 (206)
Ohio	1996						
Democratic			.234 (243)	.225 (98)		10.20 (191)	11.73 (89)
Republican			.185 (232)	.202 (98)		8.35 (160)	11.28 (87)

Figure 1: Distribution of Primary Dates Over Election Year, U.S. House (1984-2016)

* Histogram of Days Since First Primary, with seasonal cutpoints

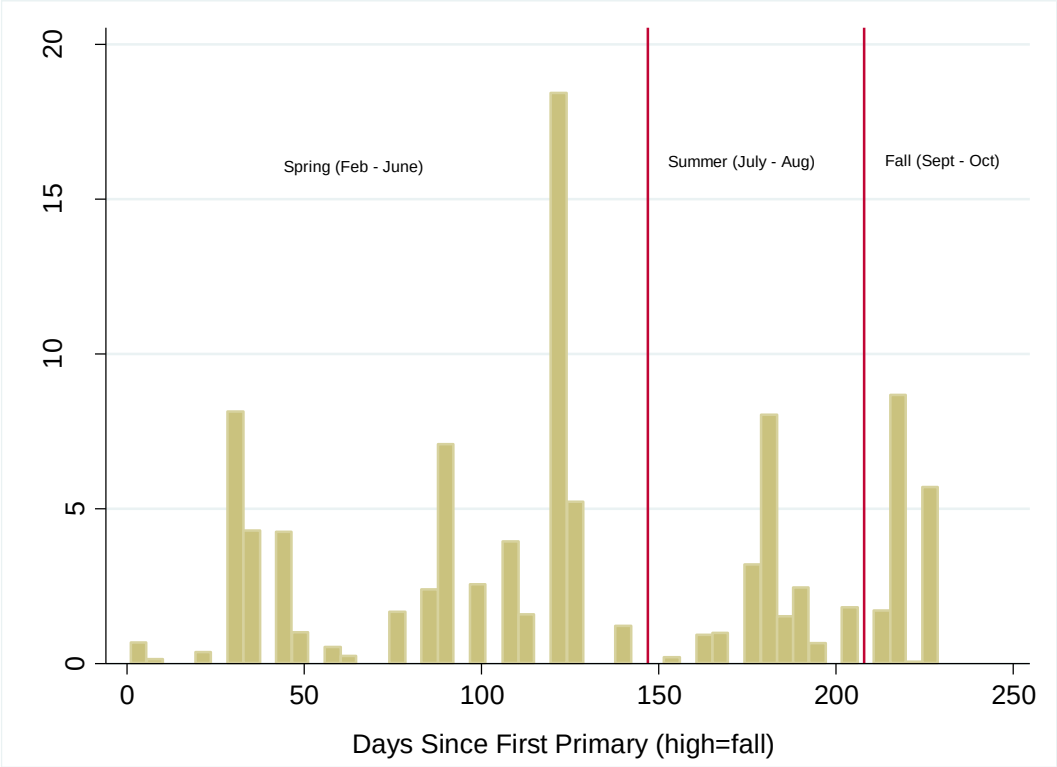


Figure 2a

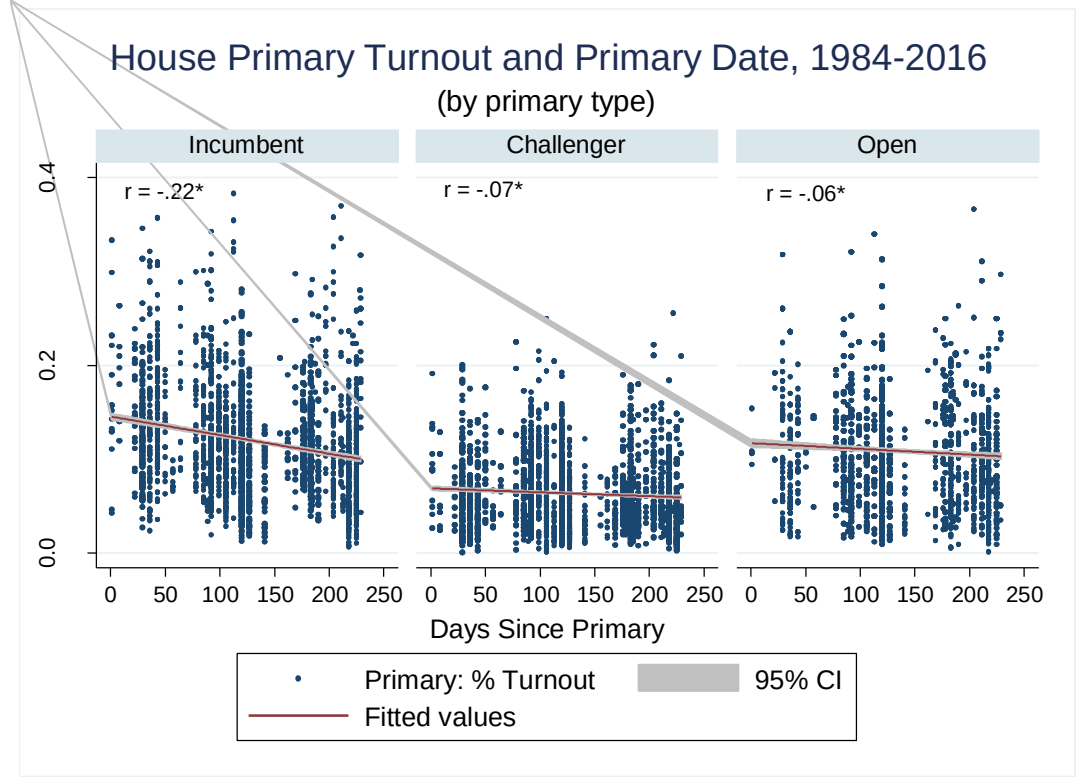


Figure 2b

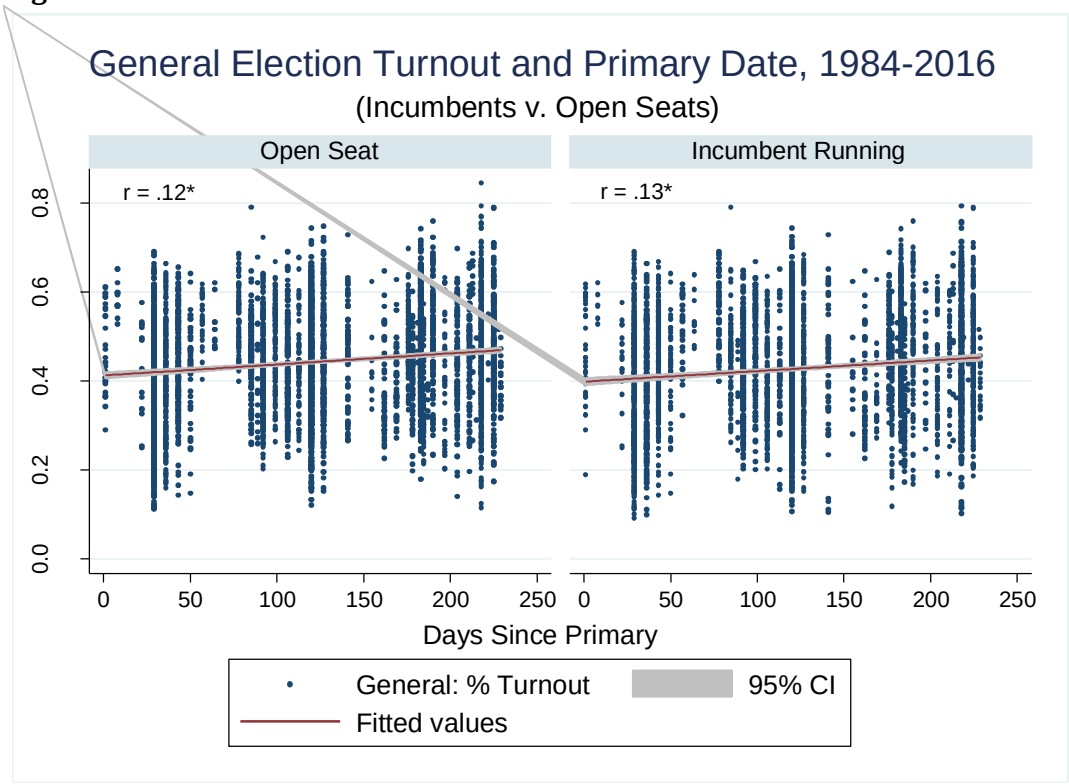


Figure 3a: Primary Turnout (%) by primary type, season, and concurrent presidential election (1984-2016)

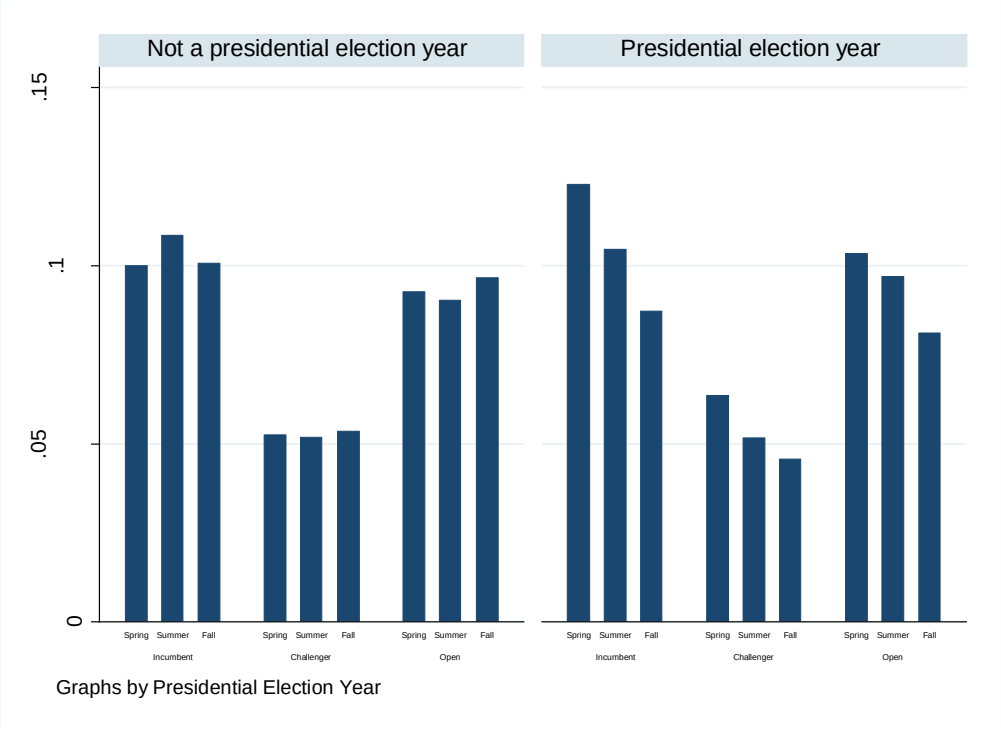


Figure 3b: Primary Turnout (%) by type, season, and concurrent presidential election (1984-2016) (New England excluded)

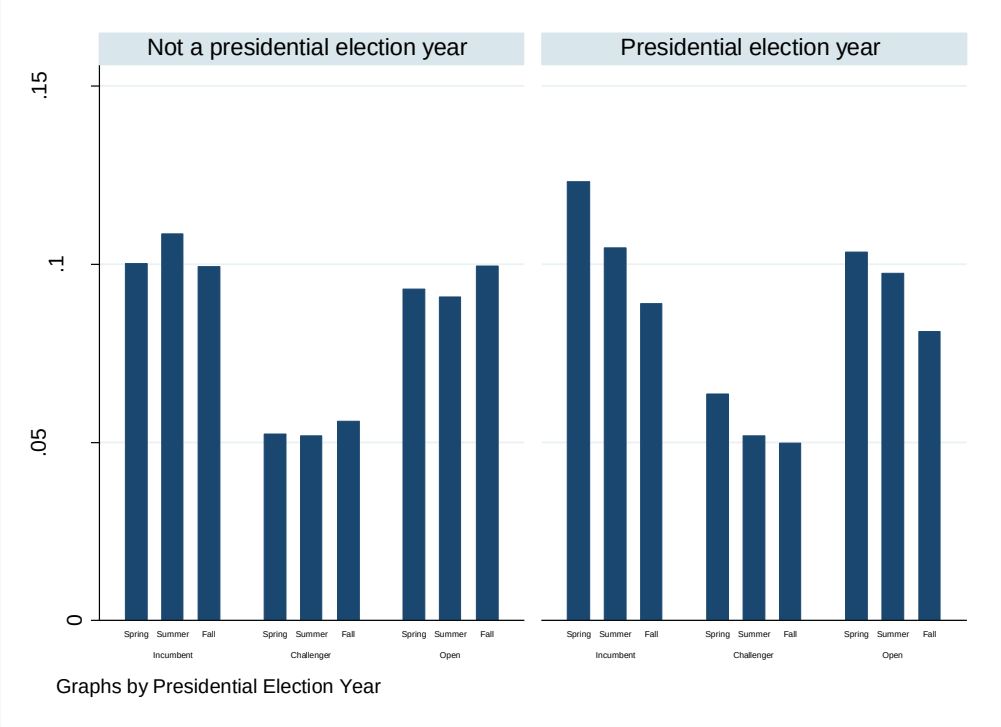


Figure 4a: Primary Turnout (%) by primary type, party, and season (1984-2016)

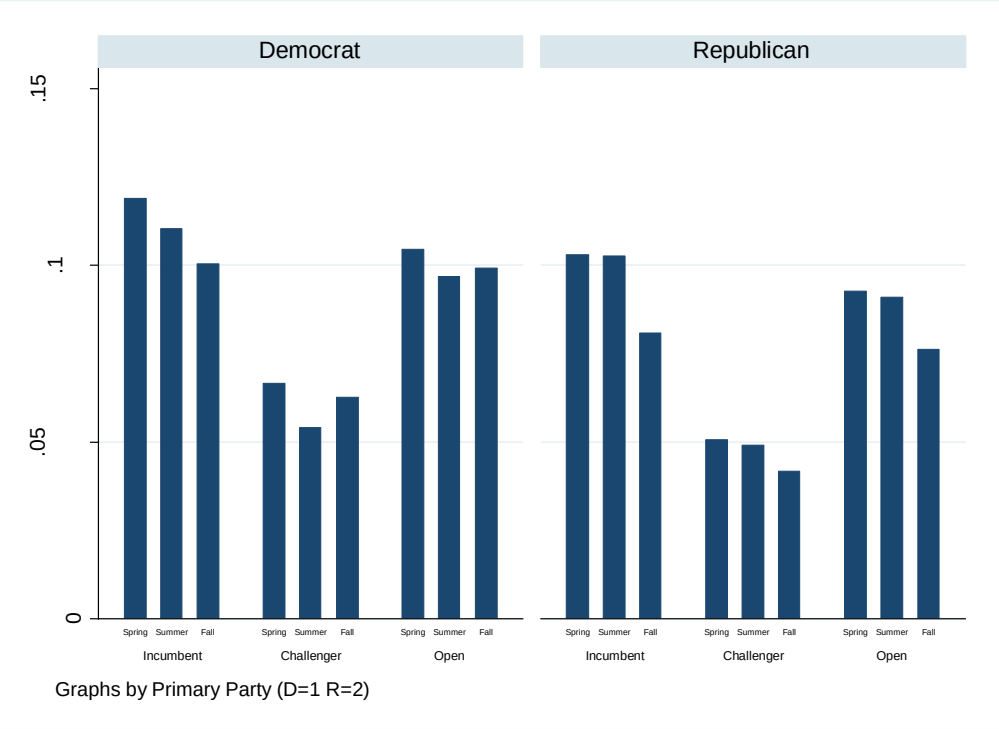


Figure 4b: Primary Turnout (%) by primary type, party, and season (1984-2016) (New England excluded)

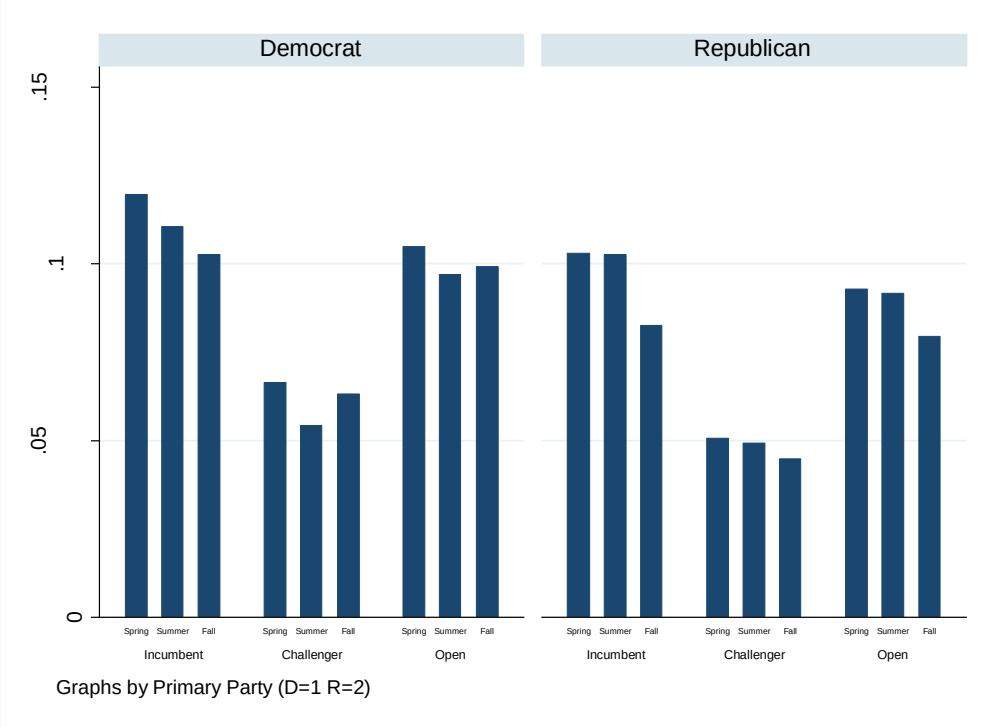
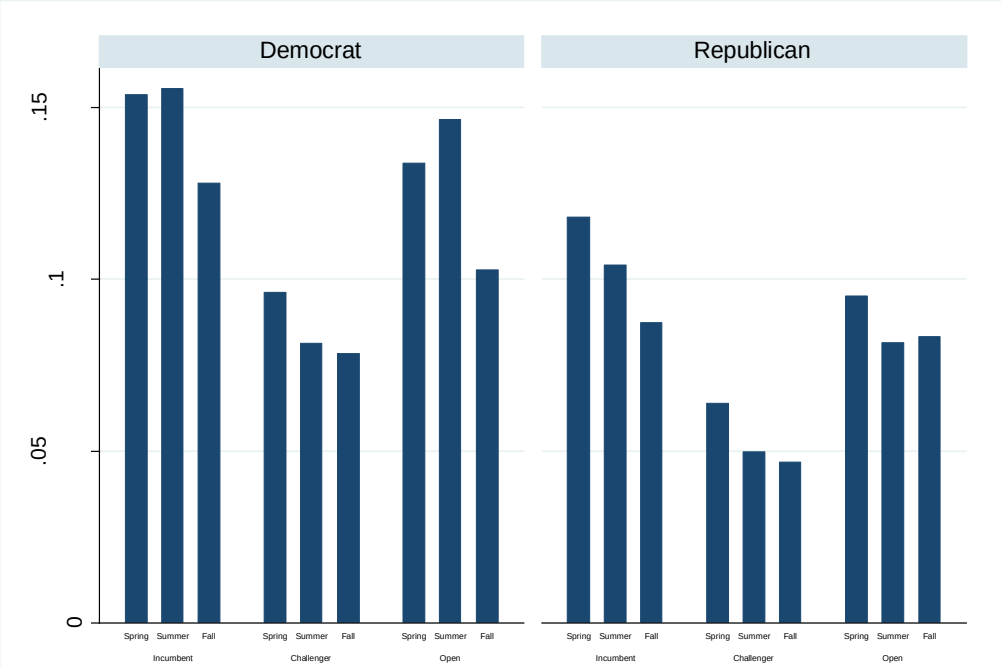
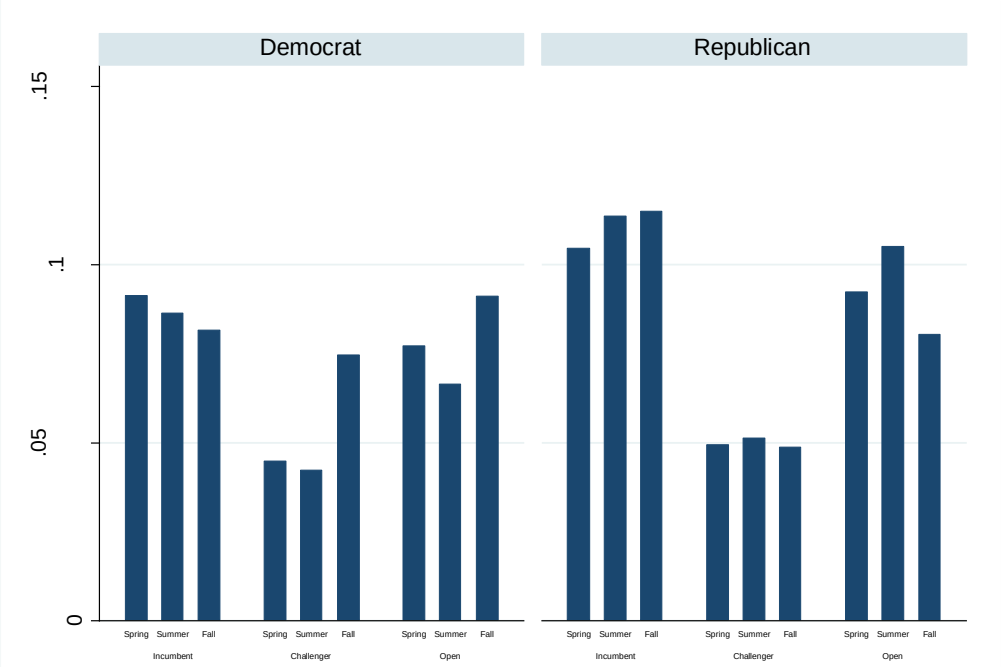


Figure 5a: Primary Turnout (%) by primary type, party, and season (1980s)



Graphs by Primary Party (D=1 R=2)

Figure 5b: Primary Turnout (%) by primary type, party, and season (2010s)



Graphs by Primary Party (D=1 R=2)

Figure 6a

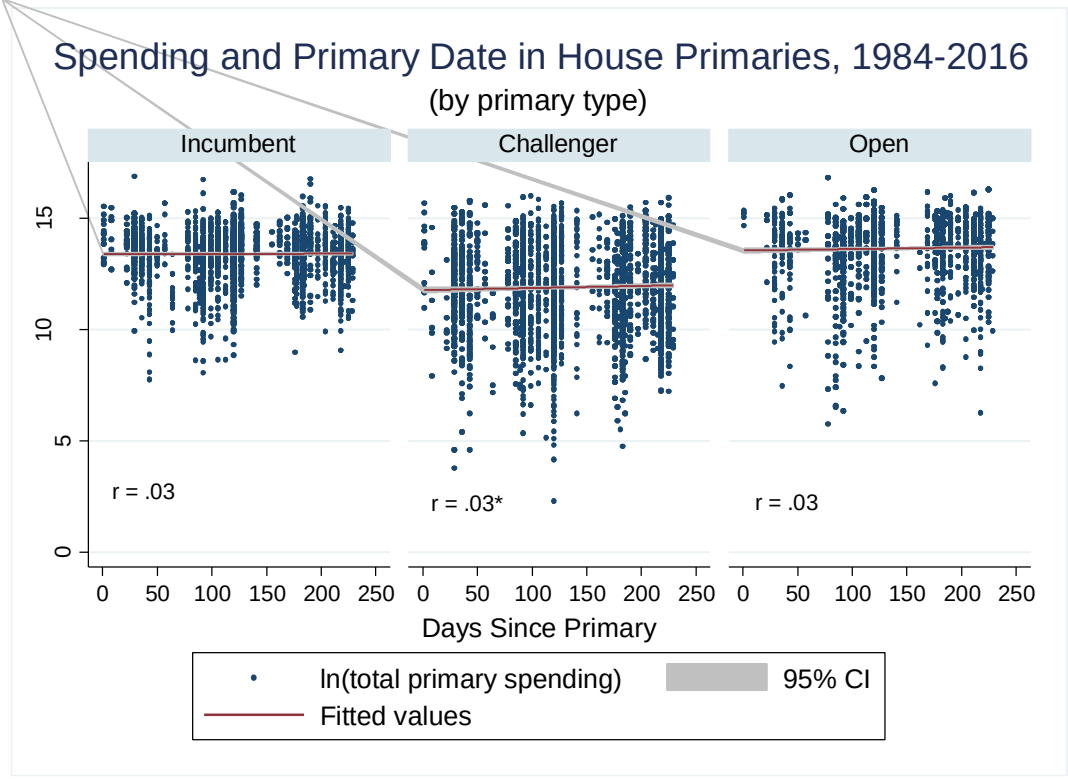


Figure 6b

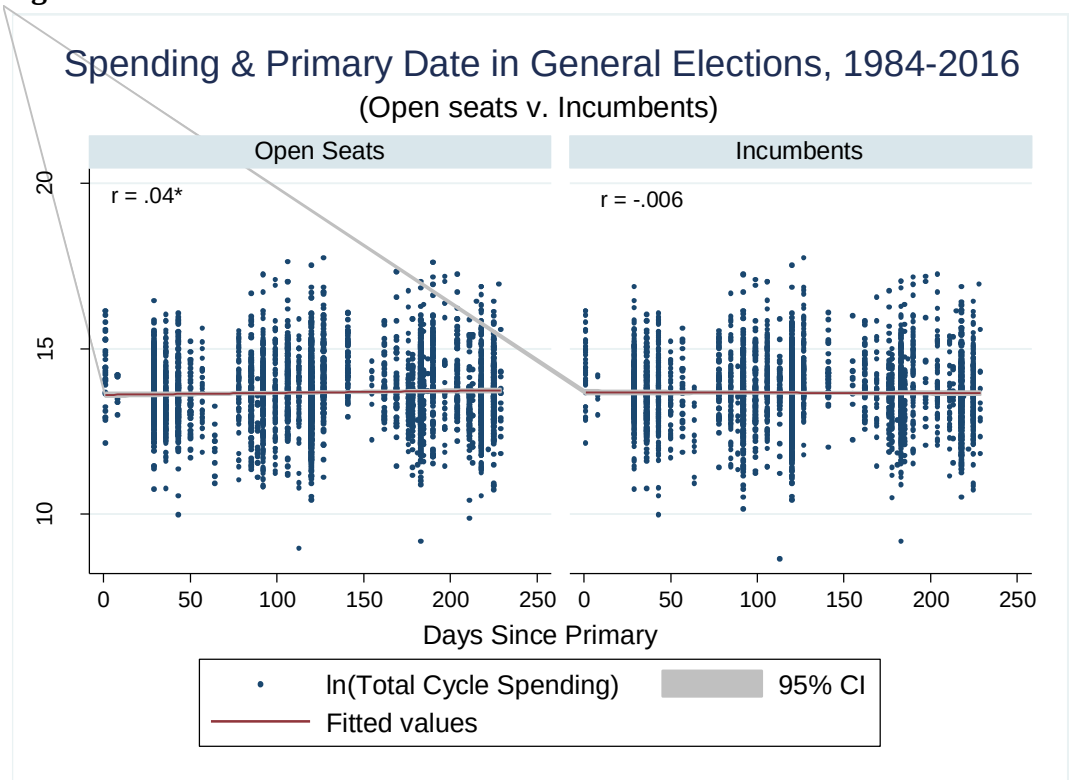


Figure 7: Mean total primary spending (thousands $) by primary type, season, and concurrent presidential election (1984-2016)

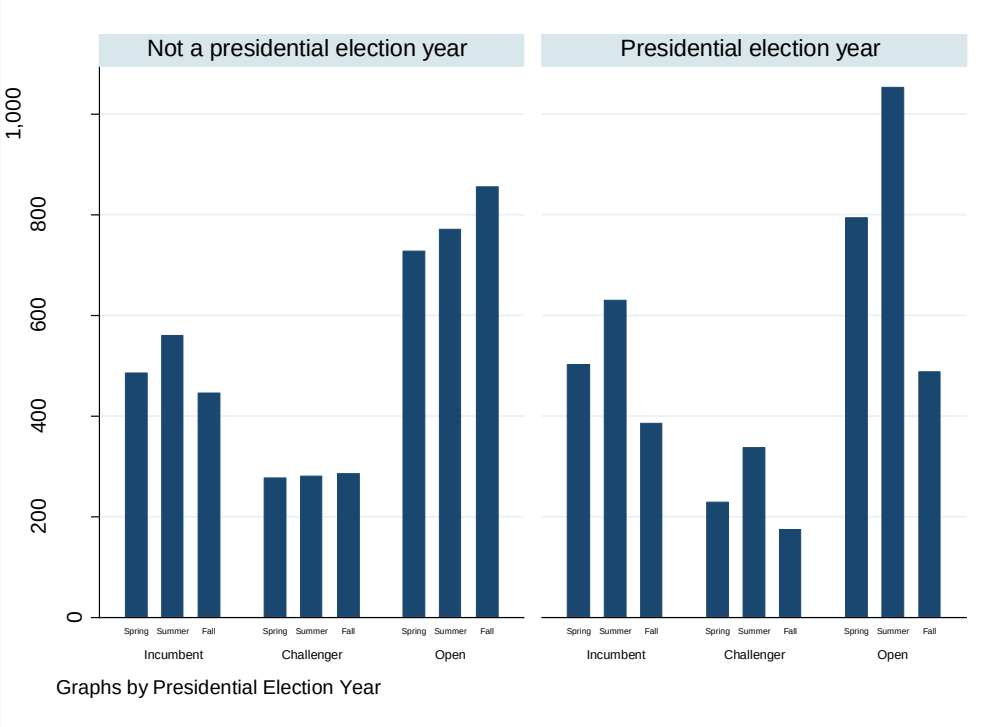


Figure 8a: Mean total primary spending (thousands $) by primary type, season, and party (1980s)

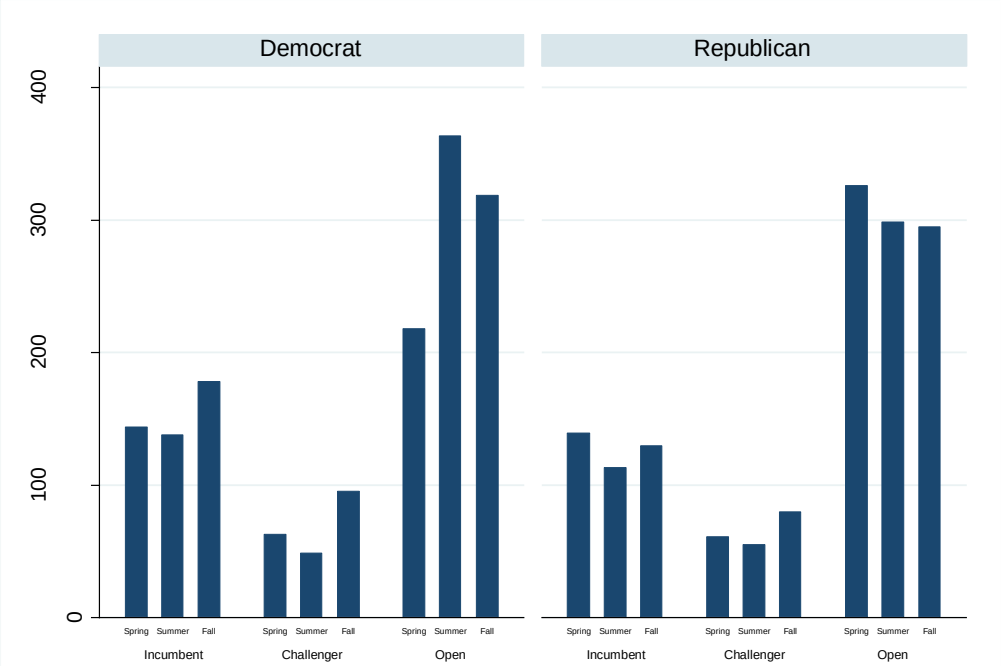


Figure 8b: Mean total primary spending (thousands $) by primary type, season, and party (2010s)

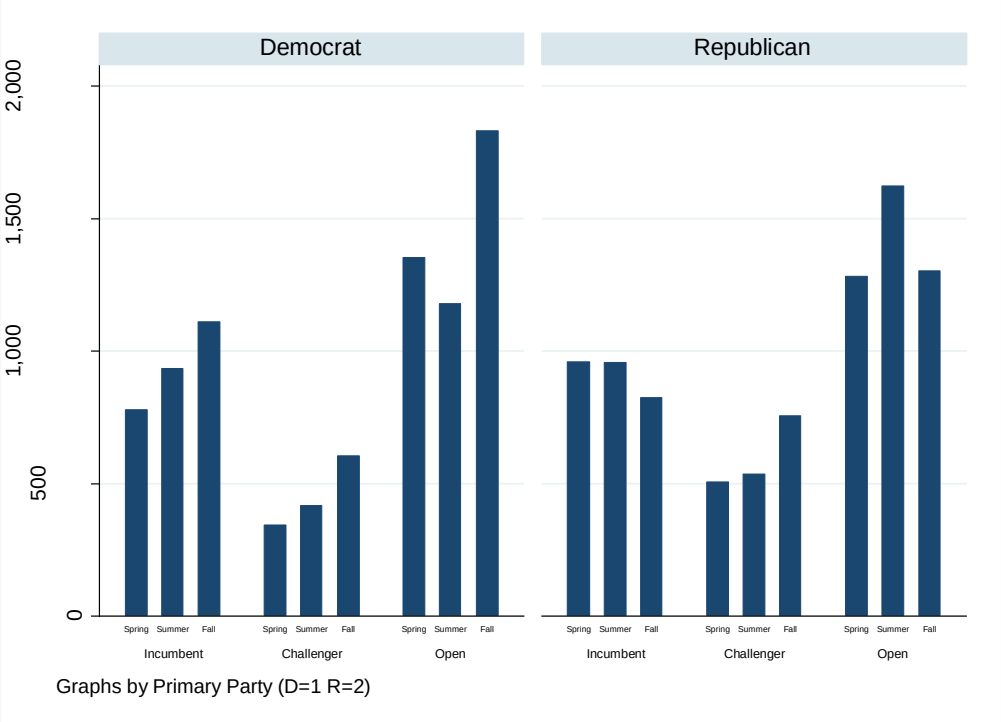


Figure 9a: Mean Total PAC Contributions in Primary, by primary type, party, and season (1984-2016)

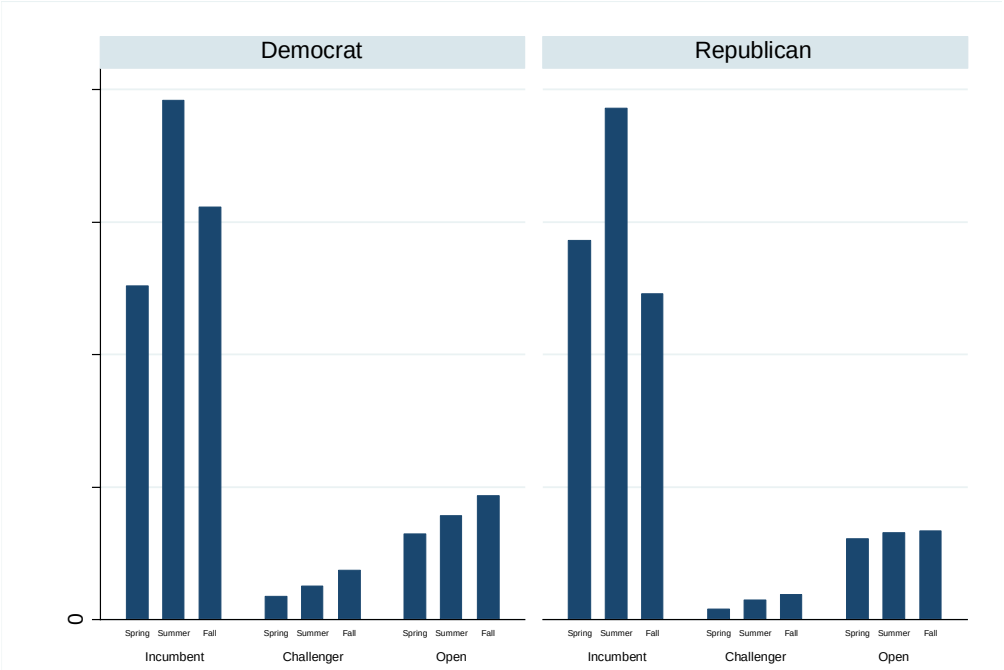


Figure 9b: Mean Total Corporate PAC Contributions in Primary, by primary type, party, and season (1984-2016)

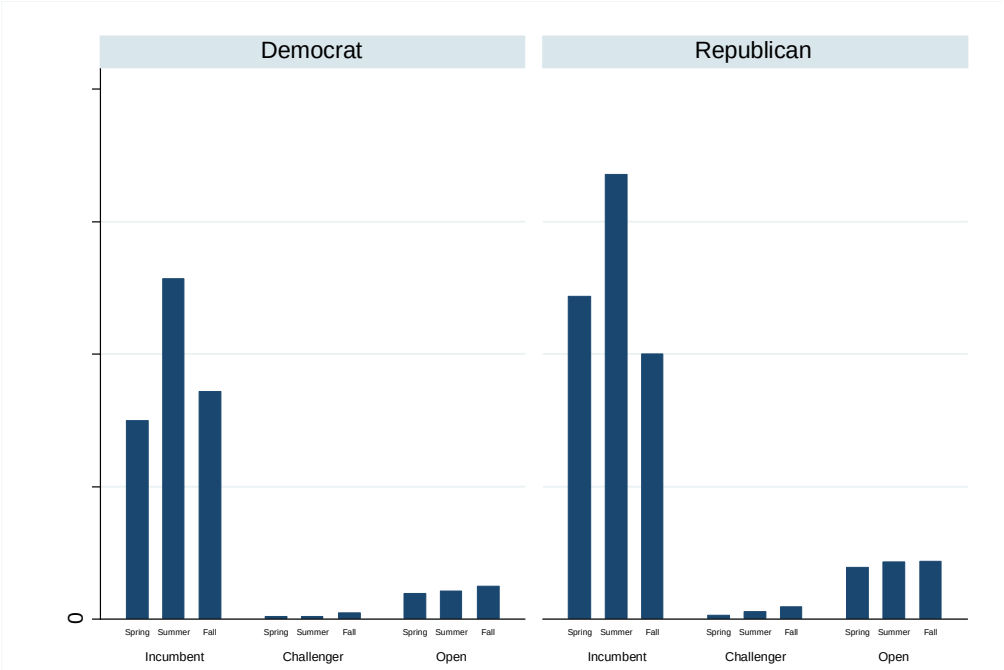


Figure 10a: Mean Total Unitemized Receipts in Primary, by primary type, party, and season (1984-2016)

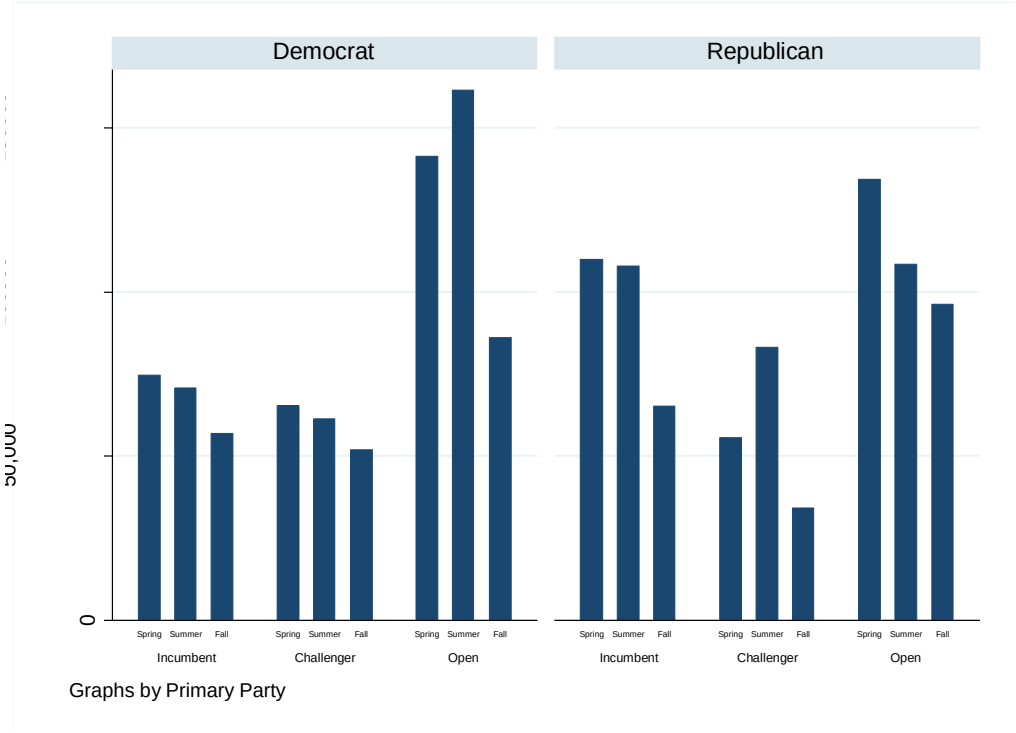


Figure 10b: Mean Total Unitemized Receipts in Primary, by primary type, party, and season (1990s)

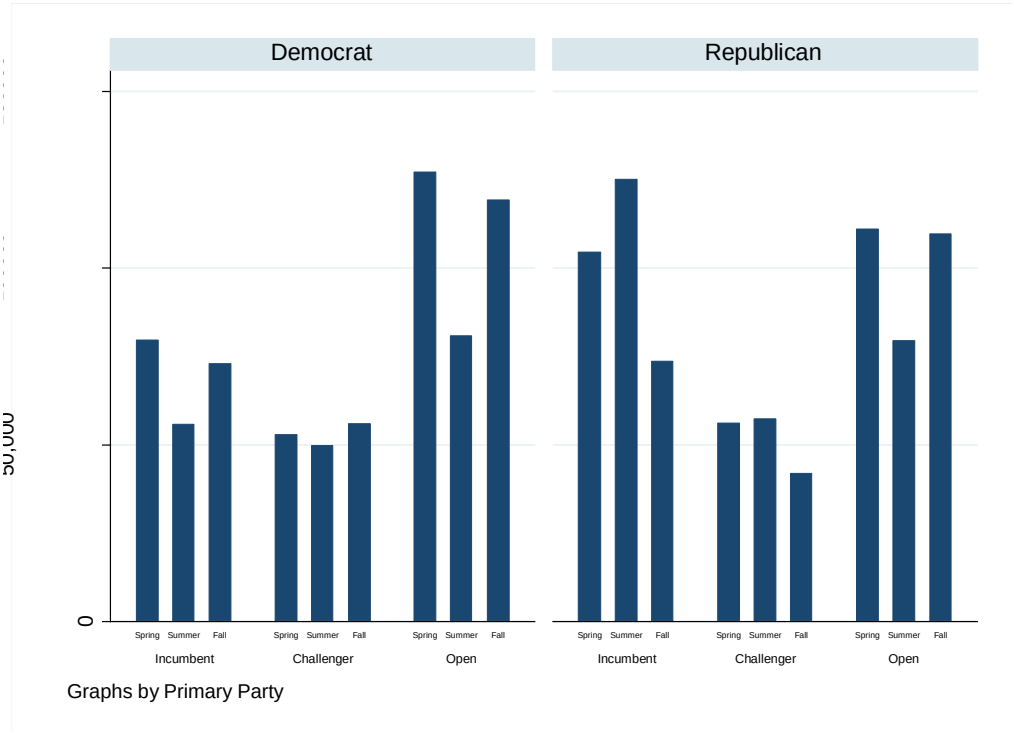


Figure 10c: Mean Total Unitemized Receipts in Primary, by primary type, party, and season (2010s)

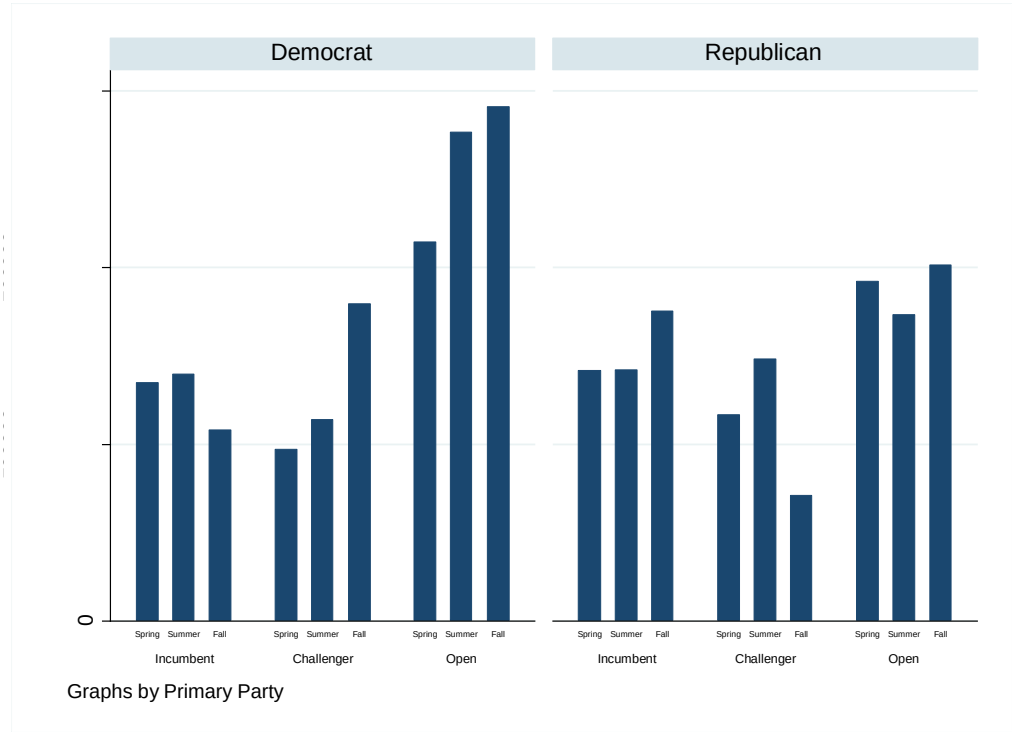


Figure 11a

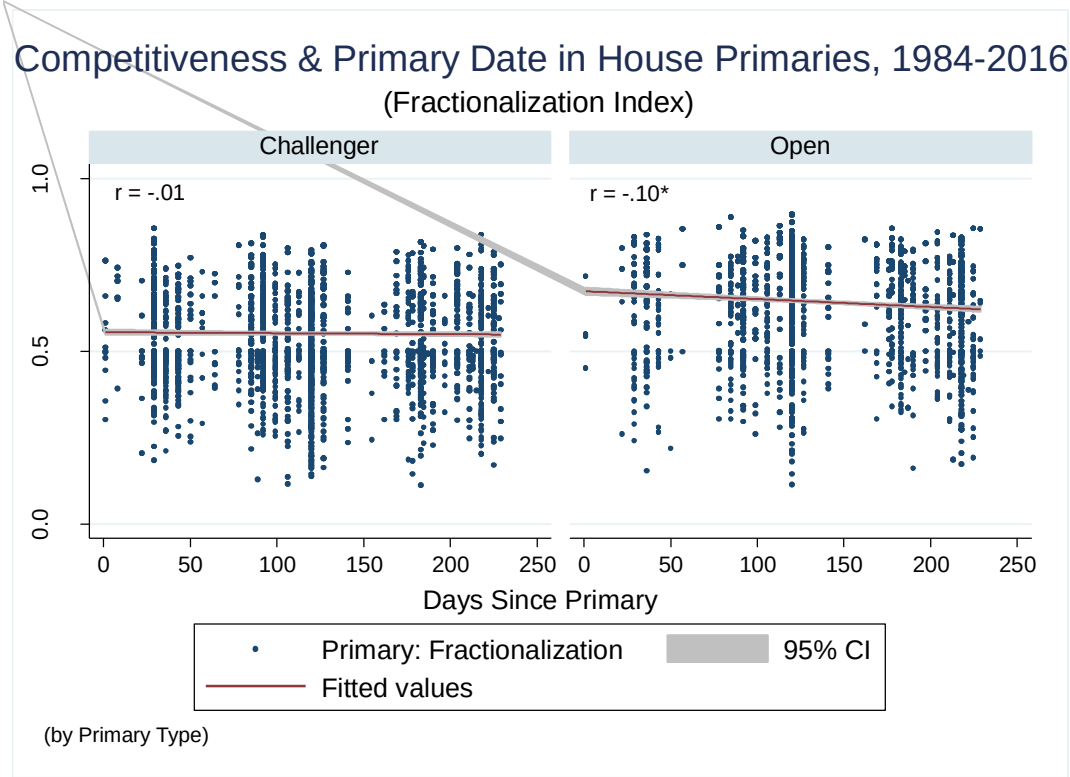


Figure 11b

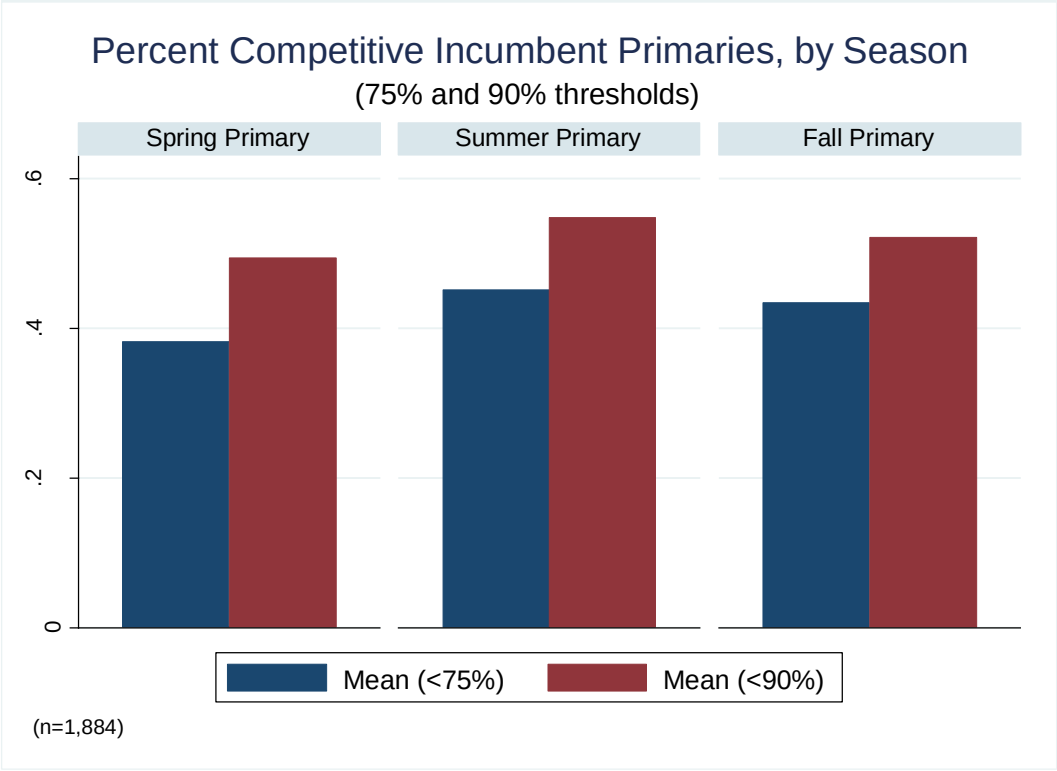


Figure 12: Competitiveness of House Primaries, by primary type, season, and concurrent presidential election (1984-2016)

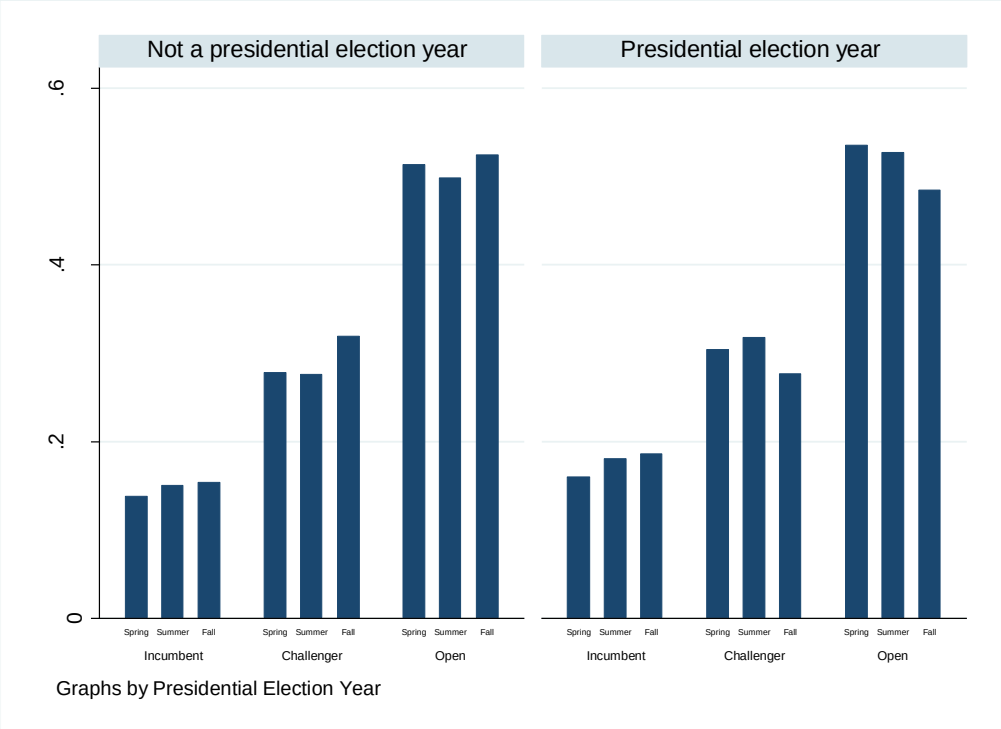


Figure 13a: Competitiveness of House Primaries, by primary type, season, and party (1980s)

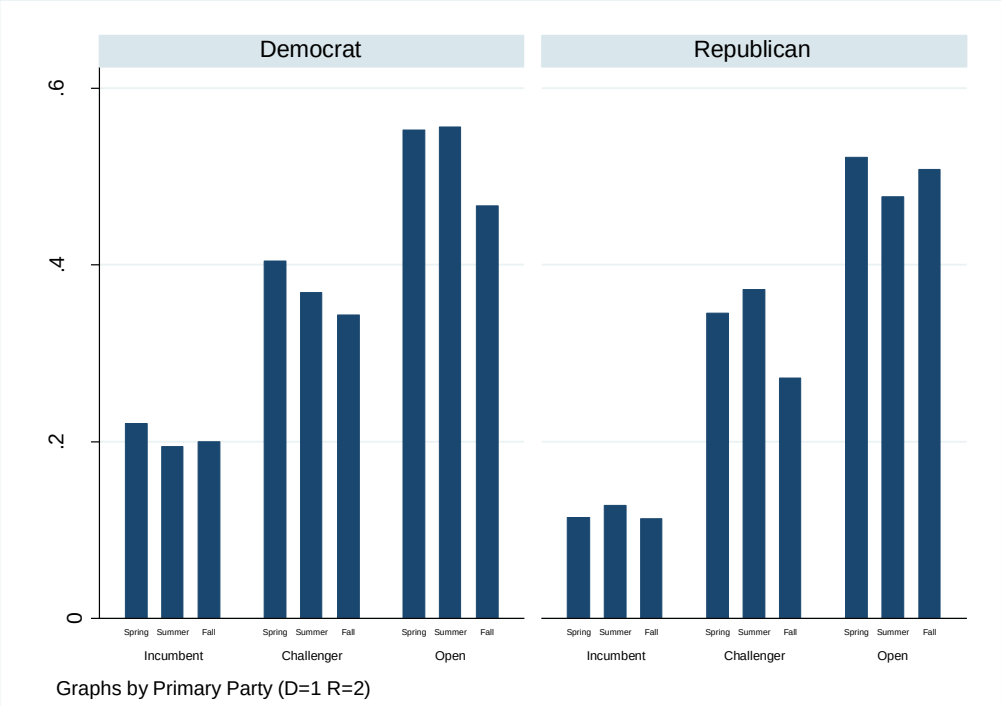


Figure 13b: Competitiveness of House Primaries, by primary type, season, and party (2010s):

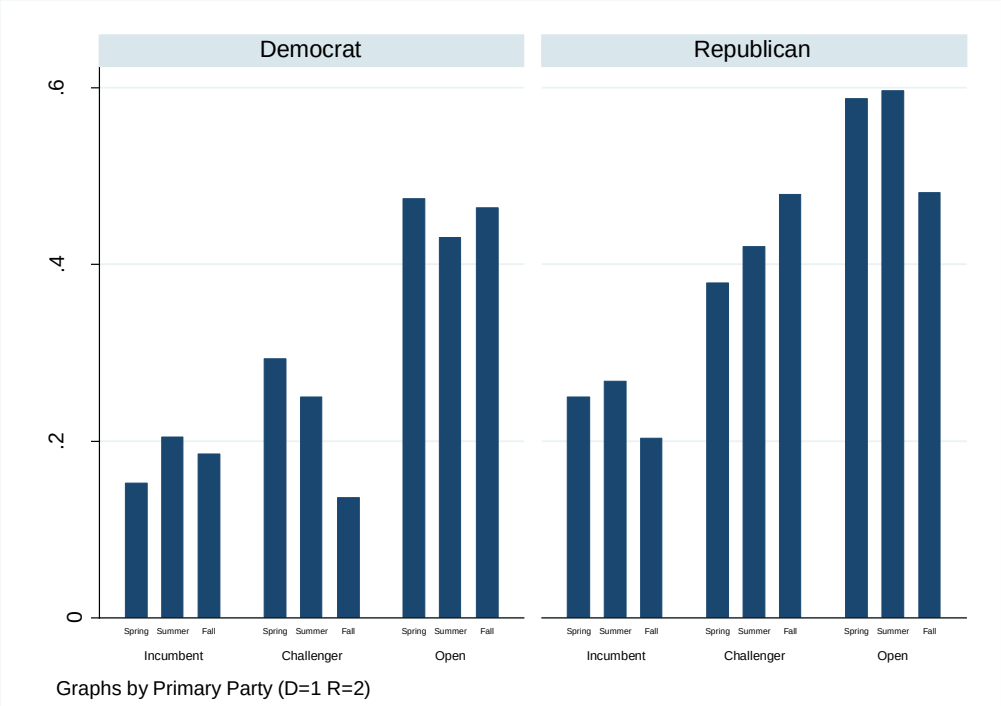


Figure 14a

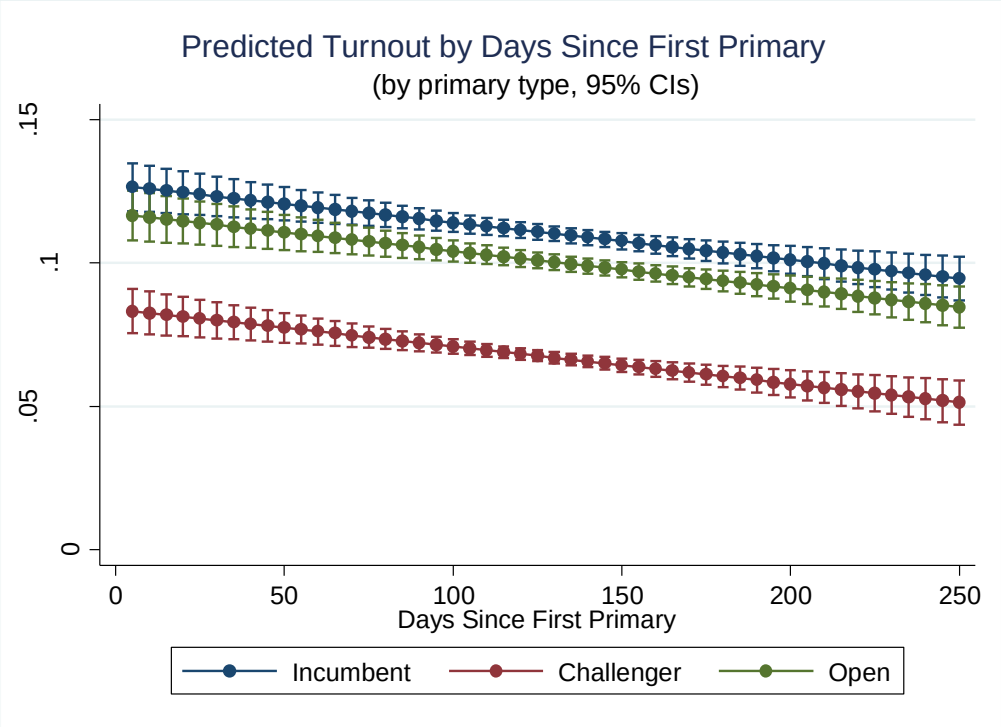


Figure 14b

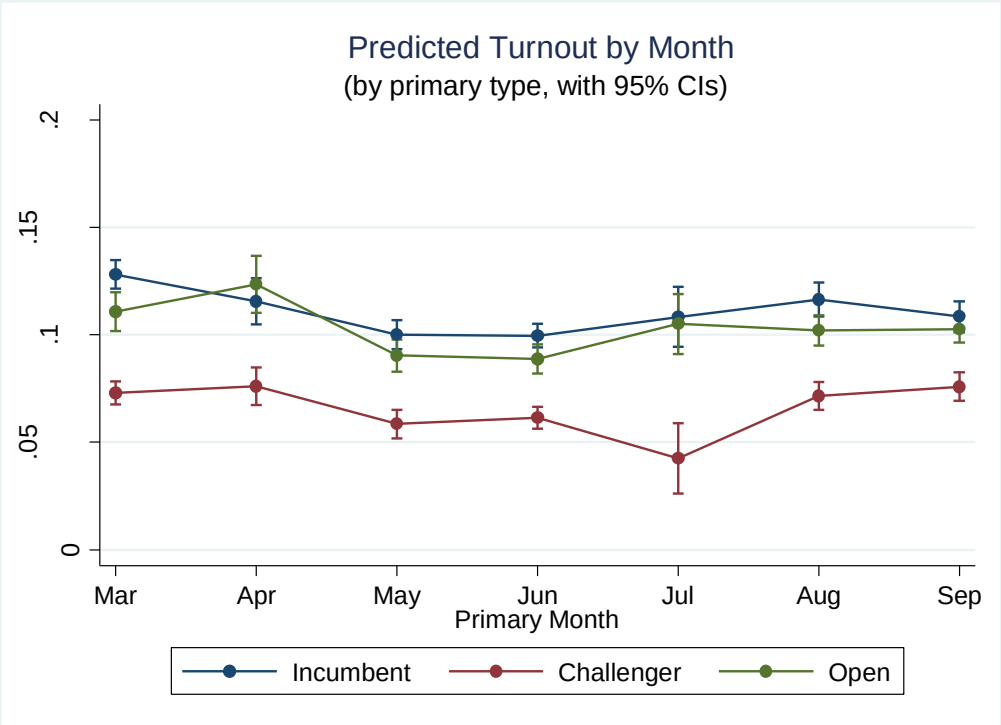


Figure 14c

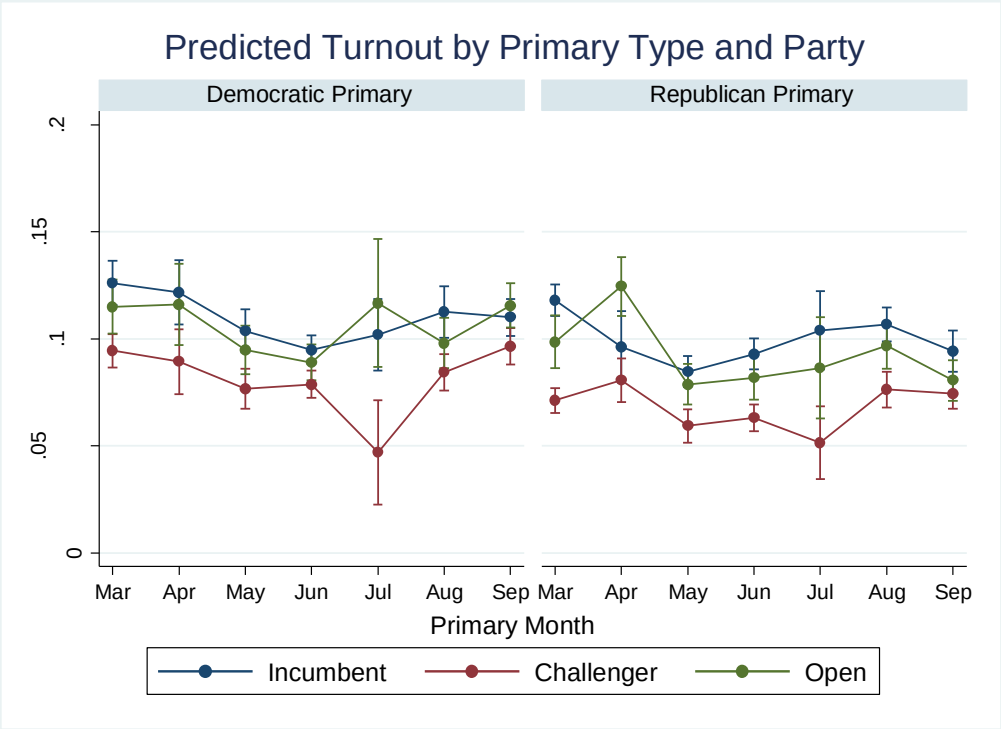


Figure 14d

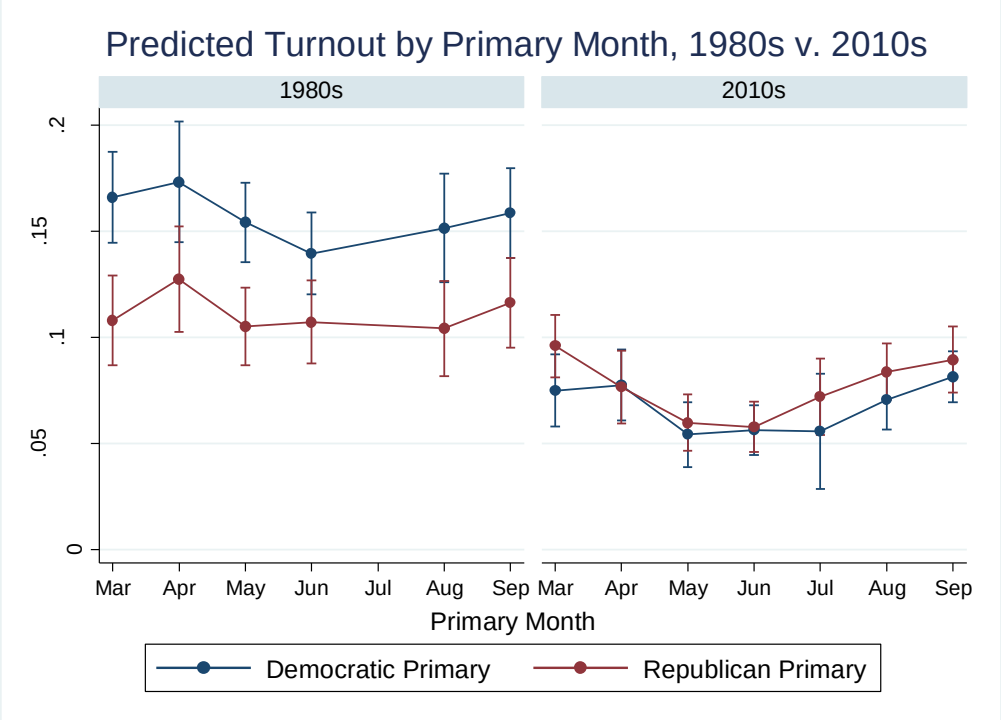


Figure 14e

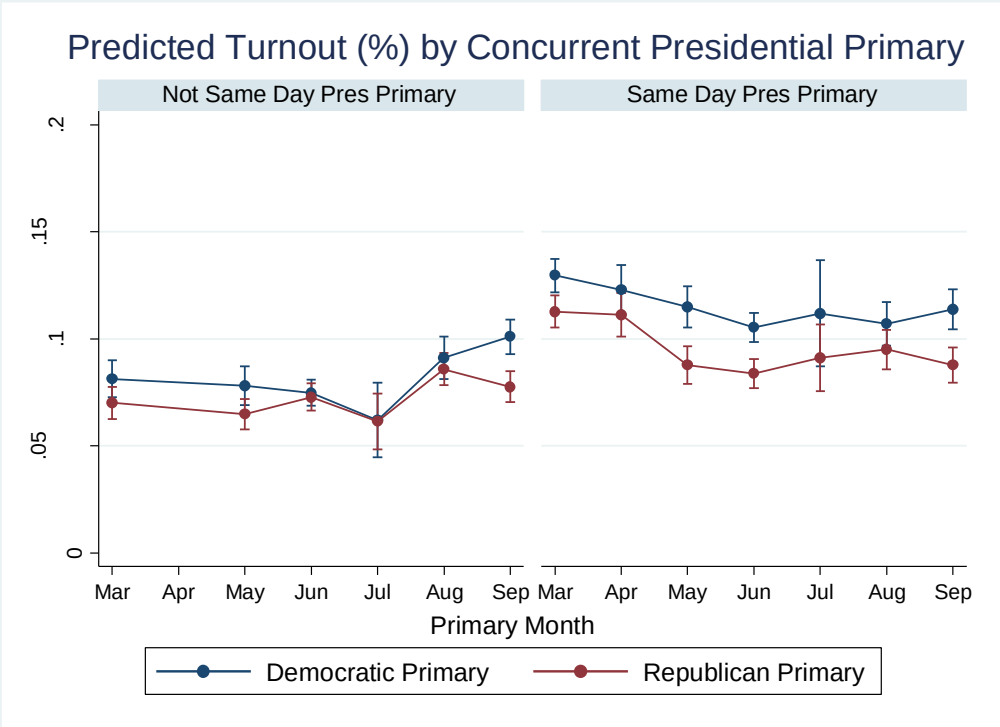

Predicted Turnout (%) by Concurrent Presidential Primary
Not Same Day Pres Primary
Same Day Pres Primary
0
.05
.1
.15
.2
Mar
Apr
May
Jun
Jul
Aug
Sep
Primary Month
Democratic Primary
Republican Primary

Figure 15a

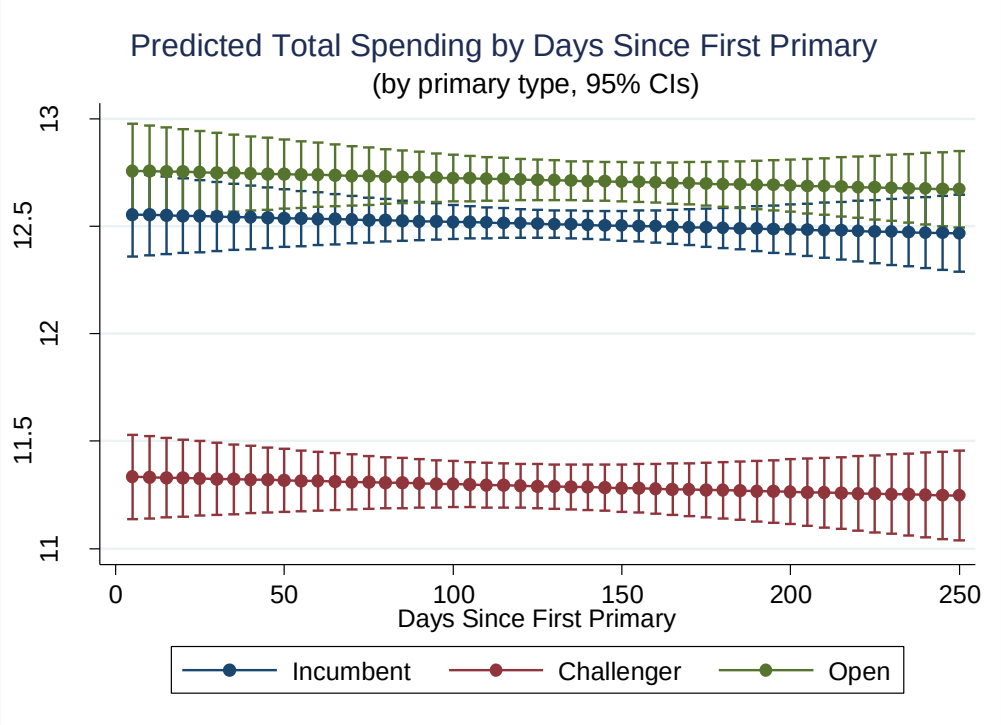


Figure 15b

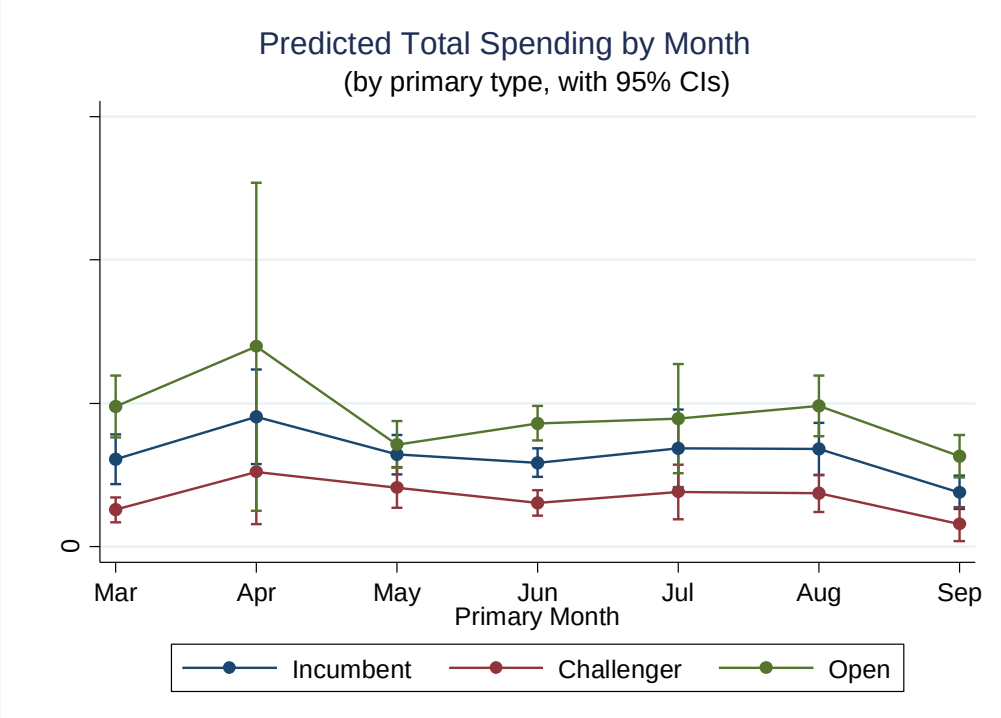


Figure 15c

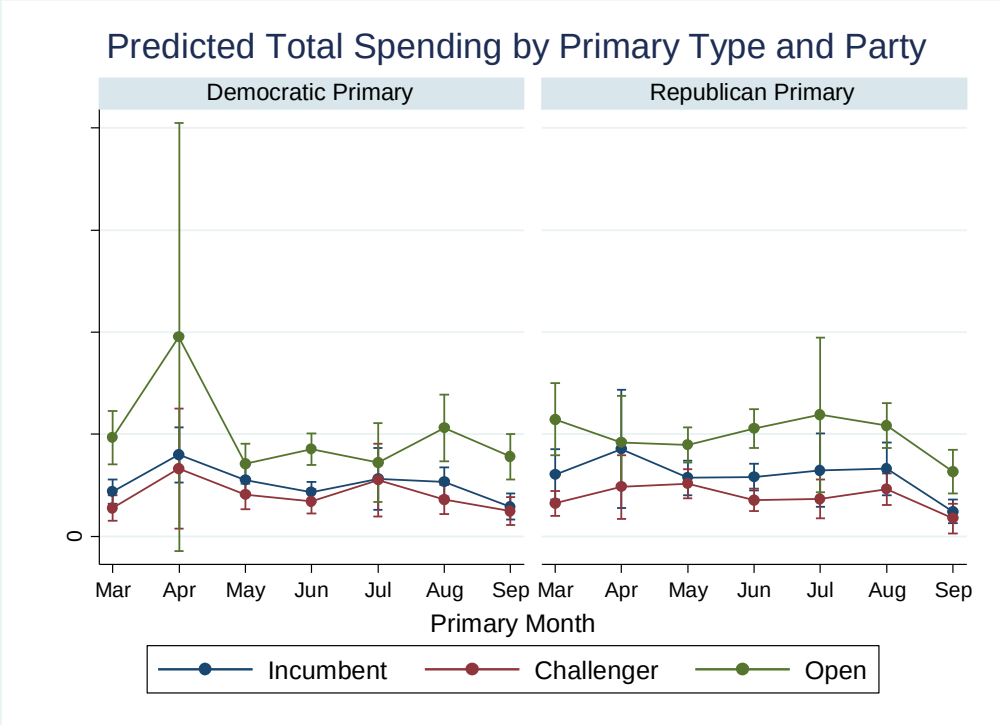


Figure 15d

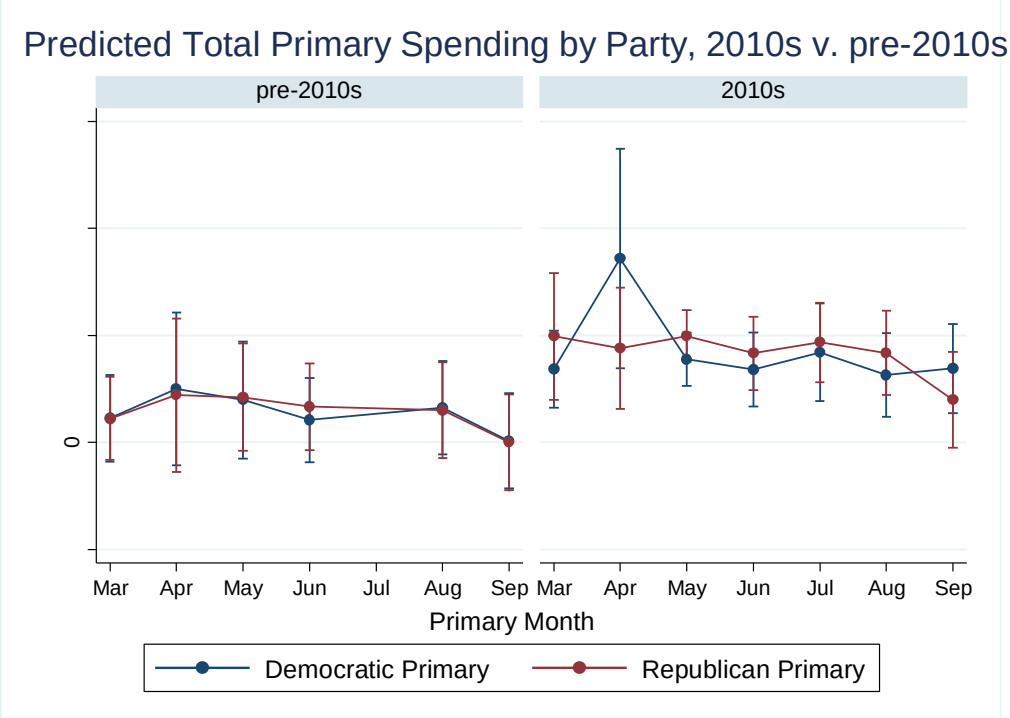


Figure 16a

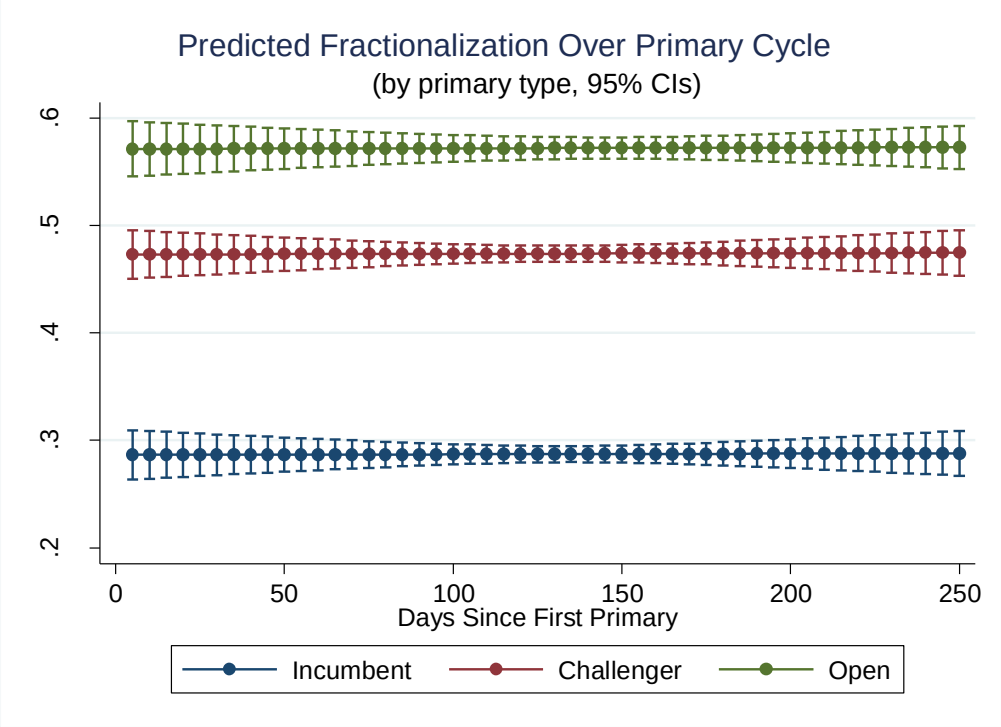


Figure 16b

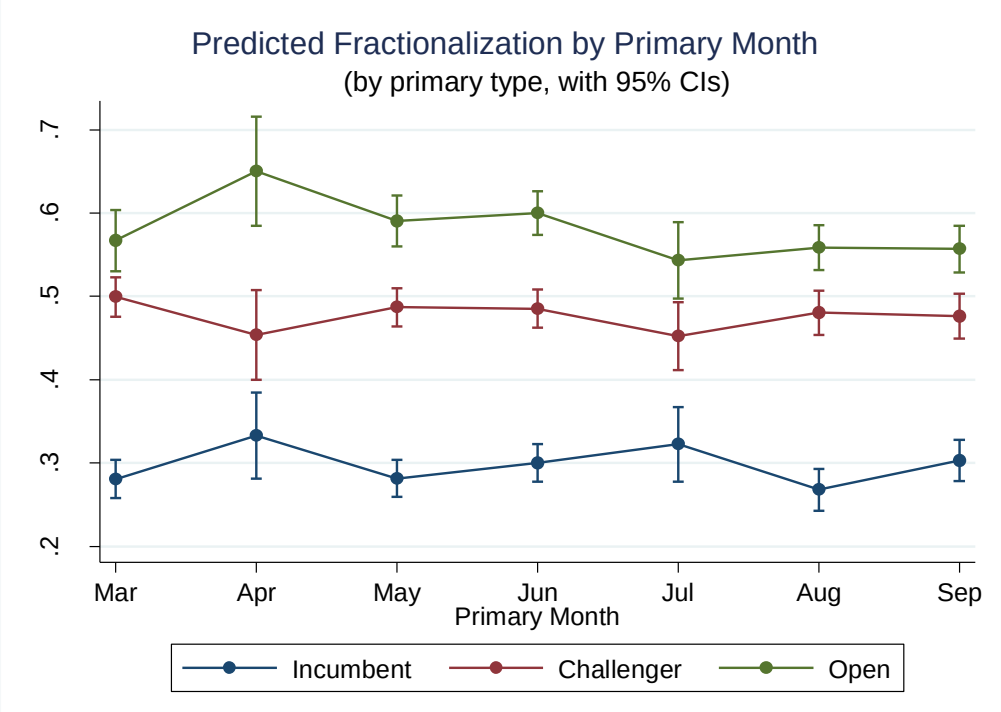


Figure 16c

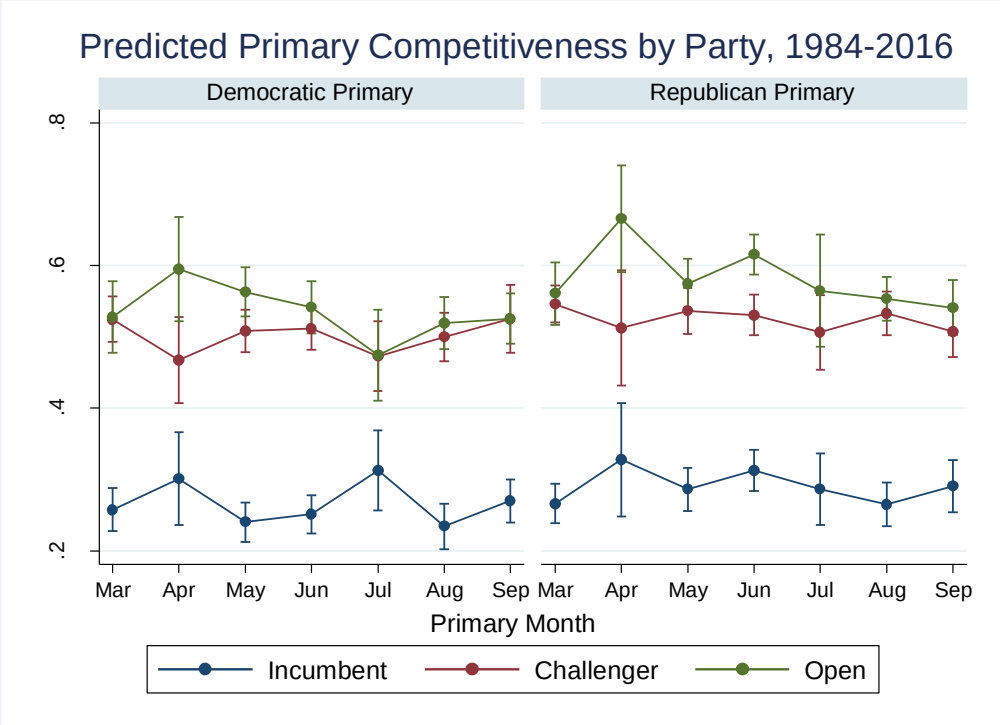


Figure 16d

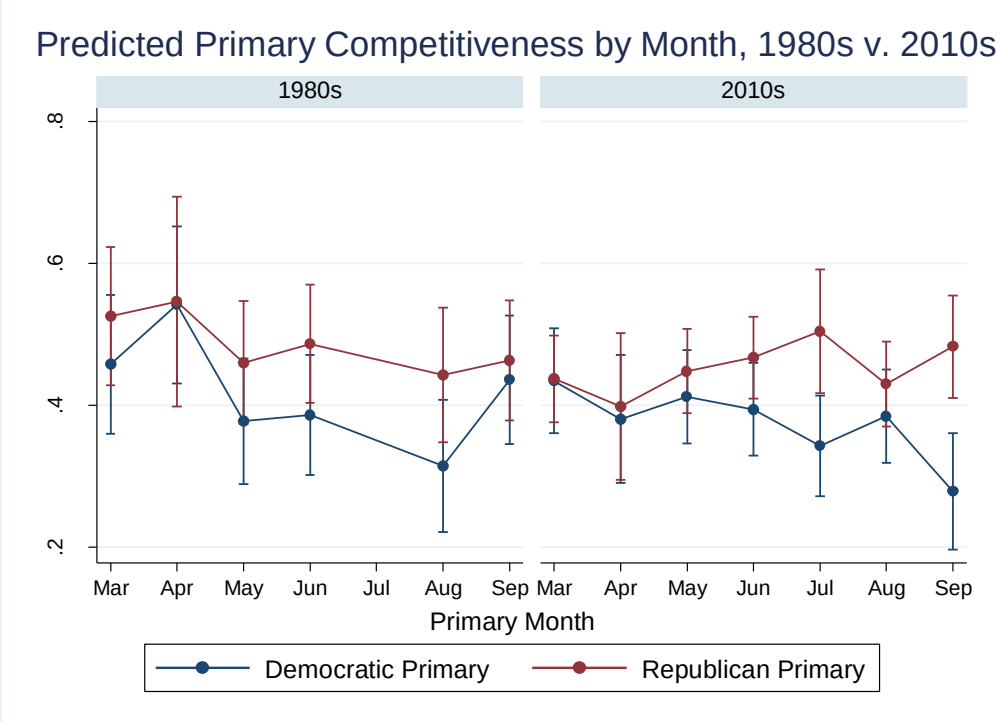


Figure 16e

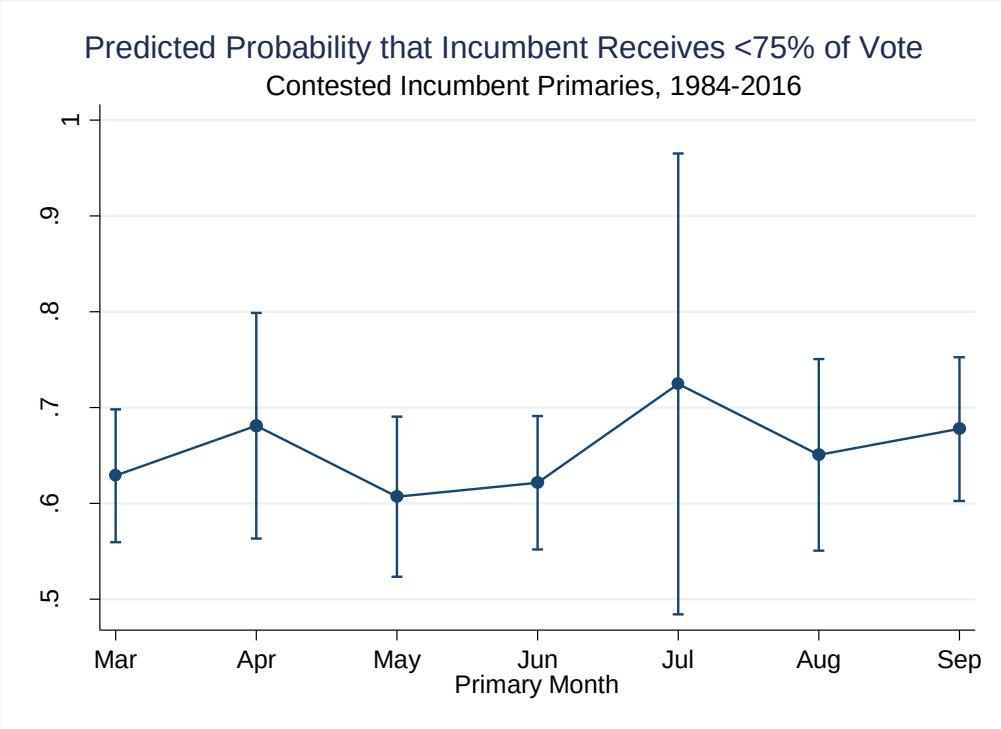

Predicted Probability that Incumbent Receives <75% of Vote
Contested Incumbent Primaries, 1984-2016
1
.9
.8
.7
.6
.5
Mar
Apr
May
Jun
Jul
Aug
Sep
Primary Month

Figure 17: Mean Number of Days Between Primary & General Elections, by Region & Year

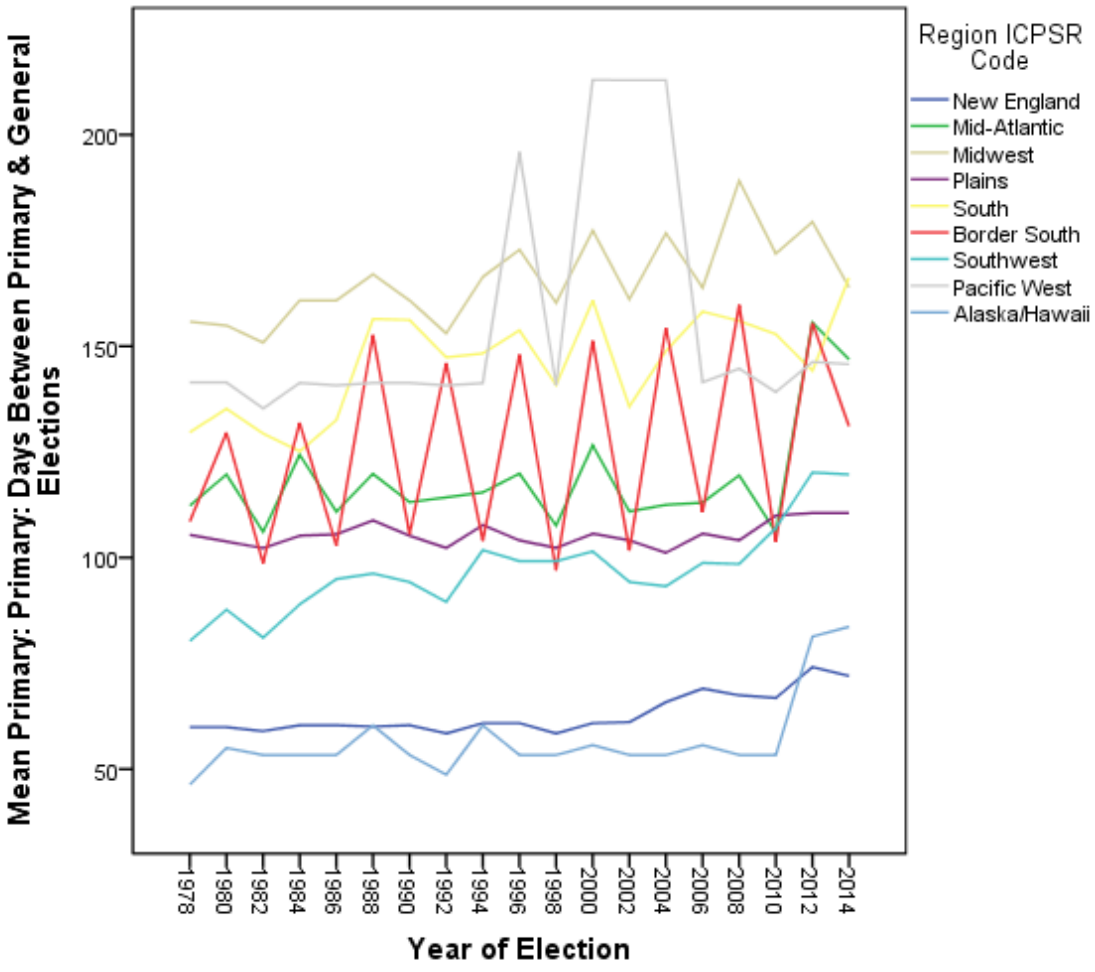
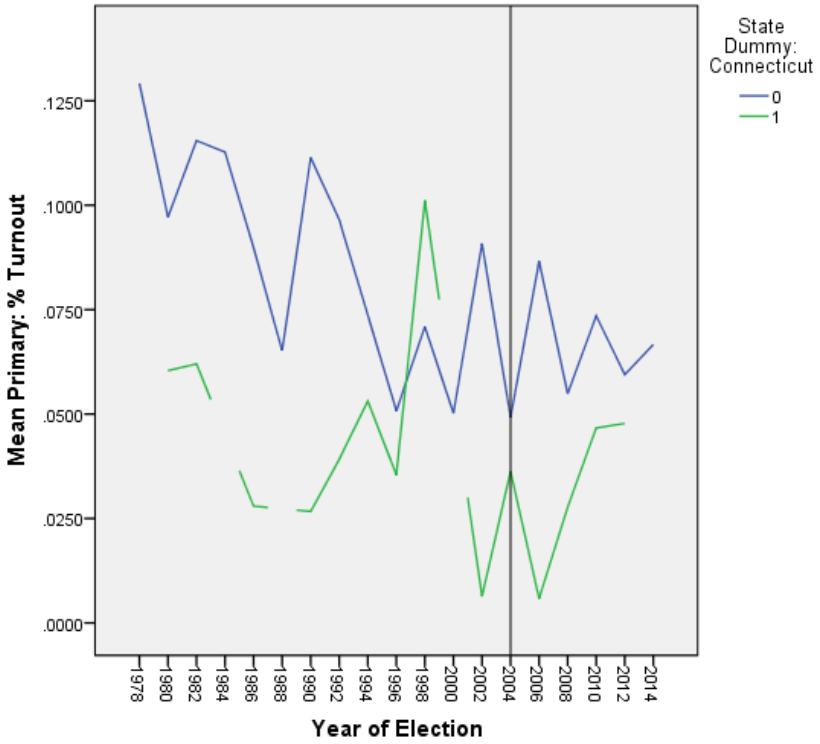


Figure 18: Voter Turnout in Selected States that Moved their Primary Dates



Connecticut
Comparison Group: Other New England states

Colorado
Comparison Group: Other Mountain states

Appendix: Notes on Dataset Construction – Clifford Vickrey, University of Connecticut

Election Variables

The first step in the creation of a comprehensive individual-level dataset of primary candidate finance was to compile the simple election returns. For years 1994 and later, federal election returns for primaries were taken from the Federal Election Commission's (FEC) election results webpage: <http://www.fec.gov/pubrec/electionresults.shtml>. For years 1992 and earlier, as well as for gubernatorial candidates, research assistants entered raw vote totals from printed matter (Scammon and McGillivray) into a cloud-computing-based web form. The House totals were cross-checked with returns collected over many years by MIT political scientist James Snyder; this triangulation ensured any remaining mistakes would be both rare and randomly distributed in the data.

General election returns were taken from the FEC, as well as the Inter-university Consortium for Political and Social Research's (ICPSR) historical election series. House returns were checked via comparison with Gary Jacobson's historical House of Representatives spending data. Senate and gubernatorial returns were checked by hand.

Once candidate names and vote totals for both primary and general election returns were in place, the next step is to add FEC identifiers. These identifiers are crucial for uniquely identifying candidates who run in multiple elections, and naturally for merging in individual-level campaign finance data. For each election cycle, the FEC identifies candidates using a nine-digit code, which contains the office sought (one digit), the ones' place of the election cycle (one digit), the postal code of the office's state (two digits, if applicable), the Congressional district (two digits, if applicable), and an iterator (three digits). FEC IDs are available back to the 1976 election cycle: on the FEC's detailed files webpage from 1979 and later, and on the FEC's public FTP server for earlier. FEC IDs were matched to year, jurisdiction, and name (the later with a strict string comparison algorithm). For unmatched cases, FEC IDs were then entered by hand.

Candidates who meet a minimal financial threshold (as of 2016, this is $5,000 in either fundraising or disbursements) must form a principal campaign committee, or PCC. The PCC, sometimes in conjunction other authorized committees, handles all donations, expenses, refunds, and committee-to-committee transactions on the candidate's behalf. Once FEC IDs are in the dataset, PCC IDs (also 9-digits) are matched from the FEC's master committee files. PCCs are necessary to merge in filing-level campaign finance information.

Unique identifiers are necessary for lagging variables, which are in turn necessary for answering questions about candidate performance over time (for instance, did an incumbent face a greater primary challenge in 2002, when her district was redrawn, than in 2000?) While it is also possible to use jurisdiction (state and Congressional district number) instead of individual as a panel variable for lagging variables, chronic district renumbering renders doing so problematic. Unfortunately, the FEC makes no special effort to ensure that candidates who run in multiple cycles receive the same ID. Whether these IDs—which are automatically generated—stay constant within candidates' electoral histories is a function of whether they run for the same

office, and list the same names and addresses in their statements of candidacy. Others are aware of this problem. The Center for Responsive Politics (CRP), in their OpenData Initiative Files, have collapsed FEC IDs into proprietary candidate identifiers. Unfortunately, these files only go back to 1990.

To solve this problem, I first collapsed disparate FEC IDs that obviously belonged to the same person automatically. I wrote a program that compared the names, jurisdictions, parties, and addresses of over one billion pairs of FEC IDs. Pairs that overwhelmingly matched on each field were collapsed into single FEC ID. The remaining, unmatched FEC IDs were then arduously matched by hand on the basis of journalistic and online sources identifying the IDs as belonging to the same person. Candidates lacking a known FEC ID (that is, did not appear in the candidate master file, or who ran for state office like governor where the FEC has no jurisdiction) were assigned 9-digit IDs beginning with the character "V," so as not to match the FEC-derived unique IDs (which begin with "H," "P," or "S").

Finance Variables

Cycle-wide campaign finance totals are readily obtainable. The FEC has made total receipts and total disbursements available in its summary files, both in fixed-width and delimited ASCII formats, all the way back to the 1980 election cycle. In addition, the CRP has total PAC receipts to, independent expenditures for, and independent expenditures against candidates by PAC industry type going back to 1990.

This is not the end of the story, however. A student of Congressional primaries runs into a simple problem with cycle-wide finance figures: the totals of primary winners, who must raise and spend money after the primary election, are incomparable with those of primary losers. The mean contested House primary from 1978 to 2016 took place 137 days before the general election, so that nominees on average campaigned for almost five months longer than their defeated opponents. If our research question is (say) whether primary timing and ballot access deadlines affect campaign finance, this discrepancy will defeat any attempt at an empirical test. Unfortunately, the FEC's online campaign summary records furnish no easy way to aggregate pre-primary transactions. Three ways around this limitation are available to the researcher.

First, one can obtain physical documentation for every primary winner. The FEC website makes available PDFs (Portable Document Format files) of reports back to the early 1990s. For older campaigns, one can visit the FEC Public Records office in Washington, D.C. and individually pull up primary figures in microfilm and old-fashioned computer printouts (the "E Index," which are candidate collections of supporting documents). Alas, obtaining every necessary document is prohibitively difficult to do for every primary winner. The scale of the manual data collection would be oppressively time-consuming, as almost 7,000 House candidates won contested primaries from 1978 to 2016.

Second, one can aggregate the millions of committee transactions. The FEC makes available records of all itemized transactions from 1979 onwards. Separate files record money transfers from individuals to committees ("INDIV"), transfers from committees to candidates ("PAS2"), and transfers from committees to other committees ("OTH"). In principle, it should be possible to code transactions by whether or not they took place before the primary, and aggregate

them by candidate accordingly. Unfortunately, documentation for these files is spotty, and the statistician responsible for their creation (Bob Biersack) is no longer at the FEC. The impressive size and scope of the files, as well as the FEC's strained technical resources (Levinthal 2013), means that they contain inaccuracies and omissions that are unamenable to correction. It is thus impossible to recreate campaign totals using them. To name just four problems, I note that, first, itemized disbursements are unavailable before the 2000s. Second, transaction records before the 1990s fail to include candidate loans to their committees, a major source of campaign financing. Third, some political action committees (PACs) report disbursements to candidates using negative signage while others report receipts received from candidates using negative signage, such that itemized values are impossible to interpret. Finally, fourth, other researchers have reported to me that transaction types are systematically miscoded in the historical data, and overworked FEC staffers can take only a reactive (as opposed to proactive) approach to correcting these.

Furthermore, even if the records were complete, they could not capture the whole of campaign finance: campaign committees are not required to itemize individual receipts of less than $200, an omission that grows more serious as one moves backwards in time: inflation-unadjusted campaign contributions are naturally smaller in the 1980s than in the 2010s. Attempts to cross-validate transaction aggregations with known totals produced large discrepancies, as well as a frustrating sense that one was trying to create a cow from a hamburger.

Finally, third, a researcher can use less widely publicized filing data to approximate primary campaign finance. The files, located on the FEC's FTP server, are fixed-width delimited ASCII files for which data dictionaries no longer exist, but it is possible to reconstruct the latter by comparing the column values with actual filings. To recreate primary totals for each candidate and election cycle, one simply has to aggregate the totals reported in each filing up to and including reports coded "12P" or "12C" (respectively, pre-primary and pre-convention reports, due twenty days before a nominating race). Back to the 1990 election cycle, this procedure is the most efficient way to generate primary receipt and spending totals.

Before 1990, however, these files are unavailable. Luckily for us, the FEC has recently launched the OpenFEC initiative in the service of making campaign finance data more accessible. Unlike previous online FEC data repositories, which are static files available for download, OpenFEC offers a modern API (application programming interface) with which researchers and developers can selectively query historical FEC records back to 1979. The API is fully documented here: <https://api.open.fec.gov/developers/>. To obtain filings, one simply obtains an API key, and sends a request to the OpenFEC server for a given committee's filings in a given election cycle. The server returns data objects in JSON (JavaScript Notation) format, which client-side scripts in a web application can read and display in its output in any number of ways.

For our purposes, we did not need to develop a web application to graph FEC data; we simply needed the data. Towards that end, I wrote a script in PHP (an easy-to-learn server-side programming language) that imported an array of House and Senate candidate PCCs in the 1980, 1982, 1984, 1986, and 1988 election cycles. The script grabbed all filings associated with each PCC and year pair, decoded the JSON objects into simple arrays, and exported these to an Excel

workbook. To get around the limits OpenFEC imposes on the rate and number of queries (at the time of the dataset's creation, this limit was 1,000 reports per hour; it has since been slackened), I used *crontab* scheduling on my personal CentOS server to run the queries 1,000 at a time at hourly intervals.

With the filings collected, I simply aggregated primary totals for receipts and disbursements. Table A1 compares the primary-level and cycle-level finance totals for major party House and Senate candidates from 1980 to 2014. As is evident, the ratio of primary- to cycle-wide financial activity is far greater for primary winners than for primary losers: whereas a nominee must continue to campaign, a primary loser must only conduct transactions necessary for the termination of their campaign (refunds, loan repayments, and so forth). By truncating campaign finance to include only those transactions up to the final pre-primary report, we accomplish the task of correcting for this difference between nominee and non-nominee totals.

Table A1. Primary Campaign Finance, 1980 to 2014*

	Mean $ Raised (Nominal)			Mean $ Spent (Nominal)		
	Primary	Cycle	% Primary	Primary	Cycle	% Primary
House						
Winners (*n* = 5,580)	$357,193	$714,057	50%	$234,559	$691,046	34%
Losers (*n* = 9,819)	$75,202	$114,123	66%	$68,237	$114,170	60%
Senate						
Winners (*n* = 703)	$2,455,075	$4,940,190	50%	$1,574,630	$4,929,282	32%
Losers (*n* = 1,769)	$345,983	$527,935	66%	$336,391	$530,715	63%

*Includes contested major party primaries only. Primary totals include all receipts and disbursements up to 20 days before the primary, such that the ratio of primary transactions to cycle transactions is approximate and likely understated.

Bibliography

Anzia, Sarah. 2014. *Timing and Turnout: How Off-Cycle Elections Favor Organized Groups*. Chicago: University of Chicago Press.

Baker, Paula. 2012. *Curbing Campaign Cash: Henry Ford, Truman Newberry, and the Politics of Progressive Reform*. Lawrence: University of Kansas Press.

Boatright, Robert G. 2013. *Getting Primaried: The Changing Politics of Congressional Primary Challenges.* Ann Arbor: University of Michigan Press.

Boatright, Robert G. 2014. *Congressional Primary Elections*. New York: Routledge.

Boatright, Robert G. 2016. "Partisan Competition and Primary Law Legislation, 2000-2014." Paper presented at the Annual Meeting of the New England Political Science Association, Newport, RI.

Boatright, Robert G., Michael J. Malbin, and Brendan Glavin. 2016. "Independent Expenditures in Congressional Primaries after *Citizens United:* Implications for Interest Groups, Incumbents, and Political Parties." *Interest Groups and Advocacy* 5 (2): 119-140.

Bonica, Adam. 2014. "Mapping the Ideological Marketplace." *American Journal of Political Science* 58 (2), 367-386.

Boyd, Richard W. 1989. "The Effects of Primaries and Statewide Races on Voter Turnout." *Journal of Politics* 51 (3): 730-739.

Bullock, Charles S., and A. Brock Smith. 1990. "Black Success in Local Runoff Elections." *Journal of Politics* 52 (4): 1205-1220.

Cain, Bruce E., and Elisabeth R. Gerber, eds. 2002. *Voting at the Political Fault Line: California's Experiment with the Blanket Primary*. Berkeley: University of California Press.

Center for Responsive Politics. *OpenData Initiative.* <https://www.opensecrets.org/resources/create/>.

Docherty, David C. 1997. *Mr. Smith Goes to Ottawa: Life in the House of Commons*. Vancouver: University of British Columbia Press.

Dominguez, Casey B.K. 2011. "Does the Party Matter? Endorsements in Congressional Primaries." *Political Research Quarterly* 64(3):534-544.

Elazar, Daniel J. 1972. *American Federalism: A View from the States*, 2nd ed. New York: Thomas Y. Crowell.

Elazar, Daniel J. 1999. *Minnesota Politics and Government*. Lincoln: University of Nebraska Press.

Galderisi, Peter F., and Marni Ezra. 2001. "Congressional Primaries in Historical and Theoretical Context." In *Congressional Primaries and the Politics of Representation*, ed. Peter F. Galderisi, Marni Ezra, and Michael Lyons. Lanham, MD: Rowman and Littlefield, pp. 11-28.

Geiser, Karl F. 1923. "Defects in the Direct Primary." *Annals of the American Academy of Political and Social Sciences* 106: 31-39.

Gerber, Elizabeth R., and Rebecca B. Morton. 1998. "Primary Election Systems and Representation." *Journal of Law, Economics, and Organization* 14 (2): 304-324.

Grossmann, Matthew, and David A. Hopkins. 2016. *Asymmetric Politics: Ideological Republicans and Group Interest Democrats*. New York: Oxford University Press.

Hassell, Hans J. G. 2013. "The Non-existent Primary-Ideology Link, or Do Open Primaries Actually Limit Party Influence in Primary Elections?" Paper presented at the Annual State Politics and Policy Conference, Iowa City, IA.

Hassell, Hans J.G. 2016. "Party Control of Party Primaries: Party Influence in Nominations for the U.S. Senate." *Journal of Politics* 78(1): 75-87.

Helderman, Rosalind. 2008. "Strategists Weigh In on Edwards/Wynn." *Washington Post* , February 15.

Hempstead, Ernest A. 1901. "The Crawford County or Direct Primary System." In *Proceedings of the Rochester Conference for Good City Government and the Seventh Annual Meeting of the National Municipal League,* ed. Clinton Rogers Woodruff. Philadelphia, PA: National Municipal League, 197-217.

Herrnson, Paul S., and James G. Gimpel. 1995. "District Conditions and Primary Divisiveness in Congressional Elections." *Political Research Quarterly* 48 (1): 117-134.

Horack, Frank E. 1923. "The Workings of the Direct Primary in Iowa, 1908-1922." *Annals of the American Academy of Political and Social Sciences* 106: 148-157.

Inter-university Consortium for Political and Social Research. 1995. *Candidate Name and Constituency Totals, 1788-1990*. <http://doi.org/10.3886/ICPSR00002.v5>.

Johnson, Gregg B., Meredith-Joy Petersheim, and Jesse T. Wasson. 2010. "Divisive Primaries and Incumbent General Election Performance: Prospects and Costs in U. S. House Races." *American Politics Research* 38 (5): 931-955.

Kanthak, Kristin, and Rebecca Morton. 2001. "The Effects of Electoral Rules on Congressional Primaries." In *Congressional Primaries and the Politics of Representation,* ed. Peter F. Galderisi, Marni Ezra, and Michael Lyons. Lanham, MD: Rowman and Littlefield, pp. 116-132.

Kaufmann, Karen M., James G. Gimpel, and Adam H. Hoffman. 2003. "A Promise Fulfilled? Open Primaries and Representation." *Journal of Politics* 65 (2): 457-476.

Kurlowski, Drew A. 2014. "Selection Before Election: The Direct Primary and Responsible Party Nomination." Ph.D. Dissertation, University of Missouri.

Lamis, Alexander P. 1984. "The Runoff Primary Controversy: Implications for Southern Politics." *PS: Political Science and Politics* 17 (4): 782-787.

Lazarus, Jeffrey. 2005. "Unintended Consequences: Anticipation of General Election Outcomes and Primary Election Divisiveness." *Legislative Studies Quarterly* 30 (3): 435-461.

Levinthal, Dave. 2013. "Another Massive Problem With U.S. Democracy: The FEC Is Broken." *The Atlantic*, 7 December. <http://www.theatlantic.com/politics/archive/2013/12/another-massive-problem-with-us-democracy-the-fec-is-broken/282404/>

Mann, Thomas E., and Norman Ornstein. 2012. *It's Even Worse than it Looks*. New York: Basic Books.

Mayhew, David R. 1986. *Placing Parties in American Politics*. Princeton, NJ: Princeton University Press.

McGhee, Eric, Seth Masket, Boris Shor, Steven Rogers, and Nolan McCarty. 2014. "A Primary Cause of Partisanship? Nomination Systems and Legislator Ideology." *American Journal of Political Science* 58 (2): 337–351.

Merriam, Charles. 1908. *Primary Elections*. Chicago: University of Chicago Press.

Meyer, Ernst Christopher. 1902. *Nominating Systems*. Madison, WI: The author.

National Municipal League. 1951. *A Model Direct Primary Election System: Report of the Committee on the Direct Primary*. New York: National Municipal League.

Norrander, Barbara, and Kerri Stephens. 2012. "Primary Type and Polarization of State Electorates." Paper presented at the Annual State Politics and Policy Conference, Houston, TX.
Ramiro, Luis. 2016. "Effects of Party Primaries on Electoral Performance: The Spanish Socialist Primaries in Local Elections." *Party Politics* 22 (1): 125-136.
Rogowski, Jon C., and Stephanie Langella. 2014. "Primary Systems and Candidate Ideology: Evidence from Federal and State Legislative Elections." *American Politics Research* 43 (4): 846-71.
Sabato, Larry. 1977. *The Democratic Party Primary in Virginia*. Charlottesville: University of Virginia Press.
Sabella, Jonathan D. 2009. "Primary Rules: Institutional Effects on U. S. House Primary Elections." Ph.D. Dissertation, University of Rochester.
Scammon, Richard M., and Alice V. McGillivray, eds. Various years. *America Votes*. Washington, DC: CQ Press.
Schmitt, Carly A. 2013. *Congressional Primaries and Legislative Behavior*. Ph.D. Dissertation, University of Illinois.
United States Federal Election Commission. "Detailed Files About Candidates, Parties and Other Committees." <http://www.fec.gov/pubrec/electionresults.shtml>.
United States Federal Election Commission. "Election Results for the U.S. President, the U.S. Senate, and the U.S. House of Representatives." <http://www.fec.gov/pubrec/electionresults.shtml>.
United States General Services Administration. "OpenFEC API (Beta) Documentation." <https://api.open.fec.gov/developers/>.
Zimmerman, Joseph P. 2008. *The Government and Politics of New York State*, 2nd ed. Albany, NY: State University of New York Press.