Ideological Voting in US House Primary Elections

Jacqueline Colao

Travers Department of Political Science, UC Berkeley

08 Sep 2025

Abstract

The US House of Representatives is experiencing historic levels of polarization. One possible explanation is that as parties have become ideologically sorted, primary voters have nominated more extreme candidates to general elections. This leaves the general electorate, even in competitive districts, with a choice between two extreme candidates. However, not only has existing empirical research found limited evidence that House primary voters have contributed to polarization, there is only weak evidence that ideology factor into voting decisions in House primaries at all. I argue that this could be due to a lack of survey data on House primary voters limiting the types of analyses that can be examined. To remedy this, I ran two-wave panel surveys during the 2022 and 2024 US House primary cycles, enabling the creation of a first-of-its-kind dataset of ~8,000 primary voters from 40 primaries in 10 states. This paper uses that dataset to investigate ideological voting in US House primaries and finds robust evidence of its prominence.

Introduction

Congressional polarization is at a historic high and continues to rise (Poole and Rosenthal 2000; McCarty, Poole, and Rosenthal 2006). Given the many ways in which polarization reduces the overall effectiveness of Congress as an institution (see for example Binder 1999; Hacker 2004; Hacker and Pierson 2006; McCarty, Poole & Rosenthal 2006; McCarty 2007; Sinclair 2008; Mann and Ornstein 2012), it rightfully receives much attention from scholars. As stated by Hill and Tausanovitch (2015), "Congressional polarization may be the most prominent stylized fact of American political science, and finding its source has become one of our most important research questions" (1058). Some likely sources include the ideologically-sorting of elites (Levendusky 2009) and increased party competition over control of Congress disincentivizing compromise (Lee 2009). It is unlikely that any one factor is solely responsible, but identifying drivers of polarization is

an important step in determining what might reduce polarization and in turn, improve Congress' capacity for effective policy-making.

Primary elections have been examined by scholars as one potential source of polarization. There are two main culprits discussed in the literature: primary voters and parties (conceptualized broadly as a network of policy-demanders). The voter story posits that the primary electorate is more extreme than the median party member and thus, will nominate more extreme candidates if nominations are put to a popular vote. The party story posits that primary elections are low information elections where formal parties, interest groups, and donors can organize around their prefer candidate to use money and endorsements to steer a low knowledge primary electorate towards the party-preferred candidate. In this story, this group of "policy-demanders" tend to prefer candidates with more extreme policies.

Previous seminal work on presidential primaries has presented strong evidence of party control over primaries. However, much less attention has been paid to congressional primaries. This paper presents evidence from an original panel survey of ~8,000 US House primary voters conducted in 40 primary elections during the 2022 and 2024 election cycles. I find robust evidence that House primary voters vote based on candidate ideology, particularly among voters who learn candidate ideologies throughout the primary campaign, and that party endorsements of candidates do not mitigate this effect.

There is, to my knowledge, only one study on ideological voting in House primaries, and that study looked at four House primary elections (Bawn et al. 2019). The dearth of voter-level research on US House primaries is understandable, given the lack of survey data available at the district-level that includes enough primary voters to study. Bawn et al. (2019) echo this sentiment writing, "Scholars have produced little research on policy motivated voting in House primaries. One reason is no doubt the difficulty of obtaining appropriate survey data" (534). To fill this gap, in 2022, I conducted a two-wave panel of ~2,500 respondents in 18 House primaries across 3 states. In 2024, I, along with a team of researchers, conducted an additional two-wave panel survey of ~9,500 respondents in 22 House primaries across 9 states (Colao et al., 2025) Between the two surveys, there are ~8,000 House primary voters, making this the largest panel dataset that exists on House primary voters.

This paper uses the data collected from both years and examines the extent to which voters in US House primary elections vote for the candidate closest to them ideologically. These results provide the first substantial evidence that US House primary voters are able to vote based on ideology despite the challenging information environment that plagues district-level elections. This is a key finding in regard to polarization because if US House primary voters do not factor ideology or policy into their voting decisions, either be-

cause they are not able to differentiate intra-party candidates based on ideology or for other reasons, then US House primary voters cannot be a driver of congressional polarization. I argue that the fact that House primary voters seem to vote based on ideology, even in elections where formal party organizations endorse candidates, provides strong reason for future work to continue to examine whether primary voters play a role in polarization.

This paper will first explain why there is reason to believe primaries could be contributing to polarization, despite studies that provide evidence to the contrary. Then, I will discuss the limited research on voting behavior and party control of primaries in congressional elections. Next, I will describe the survey design and data gathered, and lastly, will present evidence of ideological voting in House primaries.

Primaries and Polarization

There are two main mechanisms that can lead to primaries contributing to polarization: elite-driven primary polarization and voter-driven primary polarization. The theory of elite-driven primary polarization stems from work on extended party networks. This work conceptualizes parties as networks of party leaders, interest groups, and activists (i.e. "policy-demanders") (Koger, Masket, and Noel 2009; Bawn et al. 2012) and finds that these "extended party networks" can exert strong influence over primary election outcomes through coordinated support of their preferred candidate (Cohen et al. 2008; Desmarais, La Raja, and Kowal 2015).[1] Party elites can exert additional influence by encouraging certain individuals to run in primaries, discouraging others, and convincing non-preferred candidates to drop out (Fox and Lawless 2010; Crowder-Meyer 2013; Hassell 2016). If, on average, extended party networks prefer more extreme candidates, their influence over election outcomes could be contributing to polarization.

The theory of voter-driven polarization stems from the notion that primary voters are a highly knowledgeable and partisan subset of the electorate that tend toward more extreme ideologies, and thus, the median voter will prefer a more extreme candidate in the primary. This can happen if candidates hold divergent positions or if they converge (Chen and Yang 2002; Oak 2006). Theories that expect convergence predict that the necessity of winning the primary election to get to the general will push all candidates towards the median voter, which in this case is towards the extreme end of the party (Woon 2018).

The polarization literature has mainly tested theories of voter-driven polarization. Historical analyses have argued that it is unlikely that voter-driven primary polarization exists because polarization has continued to rise despite an increase in primaries that allow nonpartisans to vote (McCarty, Poole, and Rosenthal 2006),

[1] Some common examples of this support are endorsements, donations, and independent expenditures.

and there is no evidence that more extreme candidates were nominated to the general election after the introduction of direct primary elections (Hirano, Ansolabehere, and Hansen 2010). Numerous contemporary analyses have found no evidence that allowing nonpartisans to vote in primary elections impacts the ideology of the candidate elected (McGhee et al. 2014; Bullock and Clinton 2011). Based on these studies, primary and polarization scholars alike believe it is unlikely that primary voters are contributing to polarization.

These findings do not rule out elite-driven primary polarization. If parties (or extended party networks) are able to exert their will in primary elections and nominate their preferred candidate, the composition of the electorate and the introduction of direct primaries would not change the types of candidates elected; the party gets their way regardless. If this is the entire story, then research on polarization and ways to mitigate it should focus on who the party-preferred candidates are and the ideological proximity of the party-preferred candidate to the general electorate. It would be unimportant to study primary voters, as their role in the election is essentially superfluous. Understandably, based on evidence that there is party influence in primaries and the lack of evidence of voter-driven primary polarization, most of the scholarly attention has shifted towards parties and interest groups in primaries and away from voters.

I argue that this shift is pre-mature. Though the studies on voter-driven polarization all provide important insights, there are two main reasons to believe they're not telling the whole story. The first is that the majority of ideological sorting of the mass electorate into Democrats and Republicans is believed to have occurred after every state introduced direct primaries for Congressional elections. Ideological sorting occurred roughly between 1970 and 2000 (Levendusky 2009; Mason 2015) and the last state to introduce mandatory direct primaries for Congressional elections was Indiana in 1976. Theoretically, primary elections are most likely to contribute to polarization when voters are ideologically sorted. In the absence of ideological sorting, there is little reason to believe that primary voters only span one side of the ideological spectrum, and due to a likely wider overall ideological distribution of primary voters, little reason to believe that the median primary voter prefers an ideologically extreme candidate. The studies showing historic trends show us that primaries are likely not one of the first catalysts of polarization, but do not rule out the possibility that primaries are a current driver of polarization.

The second reason is that opening primaries up to non-partisans does not necessarily mean that the median primary voter will become an individual who prefers a more moderate candidate. This depends on the amount of non-partisans who choose to participate in the primary, the ideologies of these non-partisans, and the ideological makeup of the partisan primary voters (Chen and Yang 2002; Oak 2006). The greater the proportion of ideological extremists in the primary electorate prior to allowing non-partisans to participate, the greater the number of non-partisans with more moderate ideologies needed in order for the median

primary voter to prefer a more moderate candidate.[2] Hill (2015) finds no evidence that primary electorates in closed and semi-closed primaries differ ideologically from primary electorates in open primaries. The research on the impact of changes to primary rules do not directly examine the impact on the ideological makeup of the primary electorate itself; rather, they look at the impact on the ideological extremity of the candidate elected and find no change. This research importantly informs us that this specific type of primary reform is unlikely to reduce polarization, but does not tell us whether other types of primary reforms could have an influence on polarization, and does not tell us whether primary voters are contributing to polarization.

Theory and Hypothesis

Studies on ideological voting in Senate primaries provide reason to believe that House primary voters may be engaging in ideological voting. Hirano et al. (2015) and Hirano and Snyder (2019) find that throughout the course of a primary campaign, Senate primary voters increase their support for the candidate that is closest to them ideologically (Hirano and Snyder 2019, 229). However, ideological voting in Senate primaries does not imply ideological voting in House primaries. One reason to believe we might see ideological voting in Senate primaries but not House primaries is the different information environments; it is possible that there is more information about candidates available to voters in Senate primaries, making it easier to determine which candidate aligns with one's ideology and enables one to vote based on ideology. The decline in local news sources (Hayes and Lawless 2021) is particularly concerning for House primaries. Unlike national or state-wide races, the local nature of House districts means House elections are only of interest to a very small number of news sources, and the low salience of Congressional primaries does not provide strong motivation for these sources to cover House primaries. As local news sources further decline, these problems magnify. Though both Senate and House primaries are low salience elections, as low turnout rates in each indicate (Hirano and Snyder 2019), being a state-wide race, Senate primaries are relevant to a much greater number of news outlets than are House primaries, potentially leading to a richer information environment and more ideological voting.

Consistent with this, the small amount of extant literature on voter knowledge in US House primaries finds low levels of candidate knowledge on election day. Ahler, Citrin, and Lenz (2016) draw this conclusion based on respondents' placements of candidates on a 7-point ideology scale from "very liberal" to "very conservative". They find no relationship between respondents' relative placements of candidates and the

[2]This also depends on the ideologies of the candidates themselves, which could be endogenous to changes in electoral institutions, making the logic of opening up primaries to non-partisans even more ambiguous.

actual extremity of these candidates; voters are unable to distinguish between more extreme and more moderate candidates within party (244). Bawn et al. (2019) find a similar inability of voters to ideological differentiate primary candidates in exit polls they conducted during the 2014 primary season. They partly attribute this lack of knowledge to their findings that newspapers provide very little differentiating policy information on House primary candidates. They also find that only a small percentage of individuals can even remember the name of the candidate they just voted for when asked in an exit poll.

These studies provide valuable insight into the difficulty House primary voters have recalling candidate information, and are an important step in building the foundation of knowledge needed to understand House primary voters. But, there are reasons to believe that different measures of voter knowledge may reveal different results. We know that voters may not interpret the terms "liberal" and "conservative" in the same way researchers do (Jefferson 2020), so relying solely on measures of knowledge that incorporate those terms may lead to different results than measures with different wording. In primaries, even respondents who understand the typical-usage of "liberal" and "conservative" may have difficulty revealing their knowledge with a 7 point ideological scale. It is possible that primary voters perceive all candidates in the primary as "very liberal" (for Democratic primaries) or "very conservative" (for Republican primaries). Although 7-point scales usually capture the level of differentiation needed for analyses, in this case, respondents may need smaller increments to work with in order for their knowledge of the ideological differences between the candidates to be captured. It is also possible that the terms "liberal" and "conservative" are not picking up on the salient intra-party divides. Asking questions that get more directly at the progressive/moderate divide in the Democratic Party and the Trump/anti-Trump divide in the Republican Party may reveal knowledge that the liberal/conservative scale cannot pick up on.

Despite the decline in local news and the studies showing primary voters' low knowledge of candidates, candidate websites can be an important source of information. Though the internet is not a problem-free source of political information (see for example, Lelkes,Sood, and Iyengar 2015; Guess, Nyhan, & Reifler 2020; Allcott et al. 2024), nearly every House primary candidate has an easily accessible website. These websites can serve to greatly inform voter if 1) the websites contain ideological, policy, or valence information and 2) voters take the time to read and process these websites. Recent work has found a wealth of ideological, policy, and valence information on candidate websites. Porter, Treul, and McDonald (2024) use information from roughly 5,000 House primary candidate campaign websites from 2018 and 2020 to examine how the descriptive identities of primary candidates impact their opponents' position-taking and issue selection. Case (2024) constructs ideological scores for House primary candidates using website positions to test whether the extremity of primary candidates impact their opponents' position-taking. Meisels (2024) also uses House

primary candidate campaign websites to develop a measure of candidate ideology.

The only study, to my knowledge, that directly investigates ideological voting in House primaries is Bawn et al. (2019). Using open-ended responses from exit polls in 2014, in three of the four primary elections they study they find no evidence that voters use ideological or policy considerations to make their voting decisions and ample evidence that voters use ideological/policy considerations in one of the four elections. It is hard to know what to conclude from these results given the small number of primary elections examined. Thus, this paper looks at ideological/policy voting across a much larger number of primary elections. The data utilized covers 40 primary elections across 10 different states and two different primary election cycles, making it the most extensive examination of ideological voting in US House primaries to date. The ~8,000 primary voters in the sample enable one to not only look at average levels of ideological voting across the entire sample, but also to see whether the extent of ideological voting tends to differ based on certain characteristics of districts. Furthermore, the dataset enables one to look at ideological voting, as well as voter knowledge of candidate ideologies.

Though the evidence collected so far on ideological voting in House primaries is mixed and one may reasonably expect ideological voting to be smaller in House primaries than Senate primaries, because the House primary electorate tends to be high in political knowledge (Sides et al. 2020) and candidate websites may mitigate the deleterious effects of the decline in local news, I expect to find evidence of ideological voting.

Survey Design

In 2022, I ran a two-wave panel survey in 18 US House primary races: 12 Democratic races and 6 Republican races. The first wave was administered 25-30 days before each primary and the second wave was administered the day after the primary election. In addition to the typical benefits of panel surveys, such as controlling for individual-level confounders and being able to measure a respondent's baseline ideology/policy positions, panel surveys are particularly useful for detecting ideological voting because we can see how a respondent's vote choice shifts throughout the course of the election. Rather than using voter knowledge to draw conclusions about the likelihood of policy voting or trying to determine how strong a post-election relationship between vote choice and ideology needs to be to conclude that voters learned who to vote for throughout the course of the election, we can see whether the relationship between vote choice and ideology is stronger come election day compared to 30 days before the election[3].

The races were selected based on the likely competitiveness of the primary and the ability to recruit respon-

[3]This design is modeled after Hirano et al. (2015) who use panel surveys in primaries to study distinctions between priming and learning, and in the process detect ideological voting in gubernatorial and senate primaries.

dents. Following Bawn et al. (2019), I consider a primary "likely competitive" if it is an open seat primary race in either a competitive district or a safe district for that party, or if it is the out-party primary in a competitive district with an incumbent. Additionally, I include races where incumbents face a high quality primary challenger. Respondents were recruited through emails sent via voter file email addresses. Only a handful of states with likely competitive primaries in 2022 have email addresses included in the voter file and of these I was granted access to the California, Florida, and New York voter files. The likely competitive primaries in these states consisted of 12 Democratic primaries and 7 Republican primaries. However, many of the Republican districts did not have enough email addresses included in the voter file for recruitment to be fruitful (likely due to being in more rural areas where most of the emails are coming from individuals turning 18 and registering to vote, as opposed to urban areas that get an influx of out-of-state transplants registering to vote in addition to those turning 18) and Republican districts where I thought recruitment would be fruitful, garnered much lower response rates than the Democratic districts. Thus, the majority of the 2022 sample consists of Democratic voters.

In 2024, I led a team of researchers that designed and implemented a four-wave panel survey to run throughout the 2024 US House primaries and general elections (Colao et al., 2025). The panel consists of a pre and post primary wave and a pre and post general election wave. This paper reports results from the pre and post primary waves.

Data was collected in 22 primary races: 10 Democratic races and 12 Republican races. Respondents were recruited for the pre-primary wave through voter file mail addresses provided by L2 and email addresses provided by L2 and the New York Board of Elections. Respondents were chosen for recruitment using a probability sampling frame based on the characteristics of the districts (ex. in safe Democratic districts registered Democrats had a higher sampling probability than in competitive districts where registered Republican respondents will be of interest come the general election.). Different monetary incentives were randomly assigned to respondents and indicated on the initial recruitment message. Individuals who completed the pre-primary survey and opted into receiving an invitation for subsequent surveys were invited to take the post-primary survey through the email address they provided on the pre-primary survey. Response rates met or exceeded researchers' expectations of 4% for mail recruitment, 1.5% for paid email recruitment, .75% for free email recruitment, and 40% for wave 2 follow-up response rates.[4] The pre-primary wave was fielded roughly 50 days before the primary election,[5] and the invitation to the post primary wave was sent

[4]"Paid email" refers to recruitment invitations that offered individuals monetary incentives to respond and "free email" invitations did not offer monetary incentives. All mail invitations offered monetary incentives.

[5]This timing was determined based on an analysis of campaign TV ads in the 2020 US House primaries that showed ads start to proliferate starting around 50 days before the primary election. The pre-election wave of the survey ideally captures voter preferences before voters gain knowledge of the candidate, which is why ad dissemination was used as a timing guide. Due to the surveys being customized based on the differentiating policy issues in each district, the pre-election wave cannot be

out the day after the election (with data collection stopping a week after this).

The races were selected based on the likely competitiveness of the primary, defined in the same way as 2022. Since data collection was not restricted to only states where voter file email addresses are available, the entire universe of primary races was available to choose from. The sample was selected to have a similar amount of Democratic and Republican races and roughly 2/3 of races with competitive general elections[6]. Additionally, we wanted races from a variety of regions in the US, while also selecting a handful of states to have multiple races in the sample in order to see if results differ between states. Lastly, we over-sampled districts in California, as the availability of voter file email addresses enabled us to send out more recruitment invitations compared to states where we could only send mail recruitment, and thus, allowed us to increase the overall size of our respondent pool.[7]

Ideological Voting

The simplest way to think about ideological voting is the relationship between respondent ideology and the ideology of the candidate the respondent voted for. The stronger a positive relationship is between the two, the more ideological voting we may conclude is occurring. There are two main issues with this simplification. The first is that respondent's may adjust their reported ideology to match their preferred candidate, rather than finding a candidate whose ideology matches the respondent's ideology (Lenz 2012). The second is that respondents' votes may be driven by non-ideological factors that happen to align with ideology, rather than being driven by the ideological alignment itself. I mitigate the first issue by using wave 1 respondent ideology in all models. Since wave 1 was conducted 30-50 days before the election and most respondents know very little about the candidates at this stage, their reported ideology is much less likely to be influenced by a preferred candidate compared to reported ideology near election day. Regarding the second issue, one can find some reassurance that the relationship reflects ideological voting if it stronger when using election-day vote choice (i.e. wave 2 vote choice) compare to vote intent 30-50 days before the election (i.e. wave 1 vote choice). This is the empirical pattern one would expect if respondents become more informed about candidate ideologies throughout the campaign and vote for the candidate they learned is closer to them ideologically. If the relationship weakens between wave 1 and wave 2, even if it still positive in wave 2, it is less likely that ideological voting is truly underpinning the relationship. However, merely

so far away from the primary date that these issues cannot be determined by the researchers. Thus, fielding the pre-election wave much earlier than 50 days before the primary was not feasible.

[6]Since only races with competitive general elections would provide informative wave 3 and 4 data, we over-sampled these races. But, we wanted to be able to have enough safe districts to test whether the primary results differ between safe and competitive districts, so we decided on a 2/3 to 1/3 breakdown.

[7]There was only on competitive primary in NY with enough email addresses to recruit respondents. This primary is included in our sample. The Florida primary occurred too late in the primary season for it to work with the four-wave panel design.

comparing the wave 1 and wave 2 relationships does not rule out the possibility that an increase in the wave 2 relationship reflects voters learning a non-ideological factor about a candidate that appeals to respondents of a certain ideological bend, but does not actually reflect ideological voting. To address this confounder, in the next section, I directly test whether individuals who learn candidate ideologies are more likely to vote for an ideologically-aligned candidate come wave 2.

The model specifications for relationship between wave 1 respondent ideology and the ideology of the respondent's wave 1 vote intent and the relationship between wave 1 respondent ideology and the ideology of respondent's wave 2 vote choice are simply:

(1) $y_{i(t-1)} = \alpha + \beta_1 x_{i(t-1)} + \epsilon_i$

(2) $y_{it} = \alpha + \beta_1 x_{i(t-1)} + \epsilon_i$

where y_{it} indicates the ideology of respondent i's vote choice on election day, $y_{i(t-1)}$ indicates the ideology of respondent i's vote intent in the pre-election wave,[8] and $x_{i(t-1)}$ indicates respondent i's wave 1 ideology.

For candidate ideology, primary candidates are grouped into two categories: relatively more liberal candidates and relatively more conservative candidates. The former are coded as 0 and the latter are coded as 1. This is done at the primary election level, meaning that a relatively more liberal candidate in a Republican primary would likely be coded as a relatively more conservative candidate in a Democratic primary. I, along with research assistants, determined which candidates fit in which categories based on research we did on candidate positions in each primary.[9] In wave 1, candidate ideology is coded as .5 when respondents do not indicate a wave 1 vote intent (but do select a vote choice in wave 2).

I measure respondent ideology in four ways (see Table 1) and run each model above for each conception of ideology. Each measure is categorized as either symbolic, a self-identification based on labels, or operational, a set of specific policy positions (Free and Cantril 1967; Stimson 2004; Popp and Rudolph 2011), and either "liberal/conservative", picking up on traditional inter-party divides, or "progressive/moderate" (specific to Democrats), picking up on intra-party divides. The first is a continuous measure of respondent self-reported ideology on a scale from extremely liberal to extremely conservative.[10] I refer to this measure as "liberal/conservative symbolic ideology". The second is an average of policy preferences based on three ANES questions: government spending on services, aid to black Americans, and immigration.[11] Respon-

[8] 30 days before the primary in 2022 and 50 days before in 2024

[9] This follows the coding process used in Hirano et al. (2015).

[10] For the 2022 this is the 70-point scale that was also used to measure respondent perceptions of candidate ideology, and for the 2024 sample it is the 100-point scale also used to measure respondent perceptions.

[11] These three questions were selected based on an analysis I conducted on the 2020 ANES data that showed that both Democrats and Republicans had intra-party variation in their responses to these questions. Since any meaningful policy voting in primaries would be based on intra-party divides, nationally or district-specific, I wanted to select questions that were likely

Table 1: Four conceptually different measures of ideology

	Symbolic	Operational
Inter-Party Divide	Liberal/Conservative Symbolic	Liberal/Conservative Operational
Intra-Party Divide	Progressive/Moderate Symbolic	Progressive/Moderate Operational

dents' answers to these three questions are averaged and normalized between 0 and 1. I refer to this measure as "liberal/conservative operational ideology".

As emphasized in their labels, these first two measures get at symbolic and operational ideology that divide both within party and across parties. However, given the possibility that these types of within party divides could occur at the national level but not the district level (i.e. Democrats within one district may all be very liberal on immigration and in another may all be very moderate on immigration), I wanted to develop measures that more directly targeted salient within-party divides that are not necessarily salient divides across party. I call the two measures of these divides for Democrats "progressive/moderate symbolic ideology" and "progressive/moderate operational ideology".[12]

To measure progressive/moderate symbolic ideology I average responses to the following two questions:

1. Do you think President Biden is . . .

a. Too moderate
b. Too liberal
c. Just right
d. Unsure

2. Do you think the Democratic party is . . .

a. Too moderate
b. Too liberal
c. Just right
d. Unsure

Respondents who choose "too moderate" are labeled with a 0, indicating that they are more liberal on this dimension, and respondents who choose "too liberal" or "just right" are labeled with a 1, indicating that

to pick up on these divides. I chose questions that did this for both Democrats and Republicans to have some policy questions that were asked to all respondents.

[12] I asked Republican respondents questions aiming to get at the pro-Trump/anti-Trump divide, but due to the low Republican response rates, I did not have enough statistical power to run those analyses. In 2024, we focused the survey on specific policy positions rather than overall ideology, and so did not ask these questions again in 2024.

they are more moderate or a "typical party member" on this dimension[13]. I then average these responses, resulting in a three point measure with values 0, 0.5, and 1.

To measure progressive/moderate symbolic ideology, I average responses to four support/oppose items: cancel all student debt, abolish ICE, reduce police boots on the ground, and provide reparations to descendants of enslaved African Americans. These items were selected based on issues being discussed at the beginning of the 2022 primary season both in Democratic primary campaigns and in liberal media. Averaging these responses creates a four point measure with values 0, 0.25, 0.5, 0.75, and 1.[14] For all measures of respondent ideology, I use wave 1 respondent positions to avoid picking up on respondents changing their positions to align with their preferred candidate (Lenz 2012).

Figure 1 shows the relationship between respondent ideology and respondent vote choice ideology for each measure of ideology. The dashed lines show the relationship with wave 1 vote intent ideology (model(1)) and the solid lines show the relationship with wave 2 vote choice ideology (model (2)). Each dot shows the mean vote choice at that point on the ideological spectrum. Because liberal–conservative symbolic ideological is based on an 100-point scale, values are plotted as binned means; dot size is proportional to the number of respondents in each bin. The figure reveals a positive relationship between respondent ideology and vote choice ideology across all measures and waves. Importantly, this is slightly stronger in wave 2 compared to wave 1, suggesting that vote choice becomes more aligned with respondent ideology over the course of the election.

[13]Respondents who select "unsure" are dropped from the analysis for "moderate/progressive symbolic ideology". Dropping respondents who answer "unsure" to both questions versus only dropping respondents who answer "unsure" to one of the questions does not meaningfully change the results.

[14]Again, dropping respondents who answer "unsure" to any of these questions versus only dropping respondents who answer "unsure" to all of these questions does not meaningfully change the results.

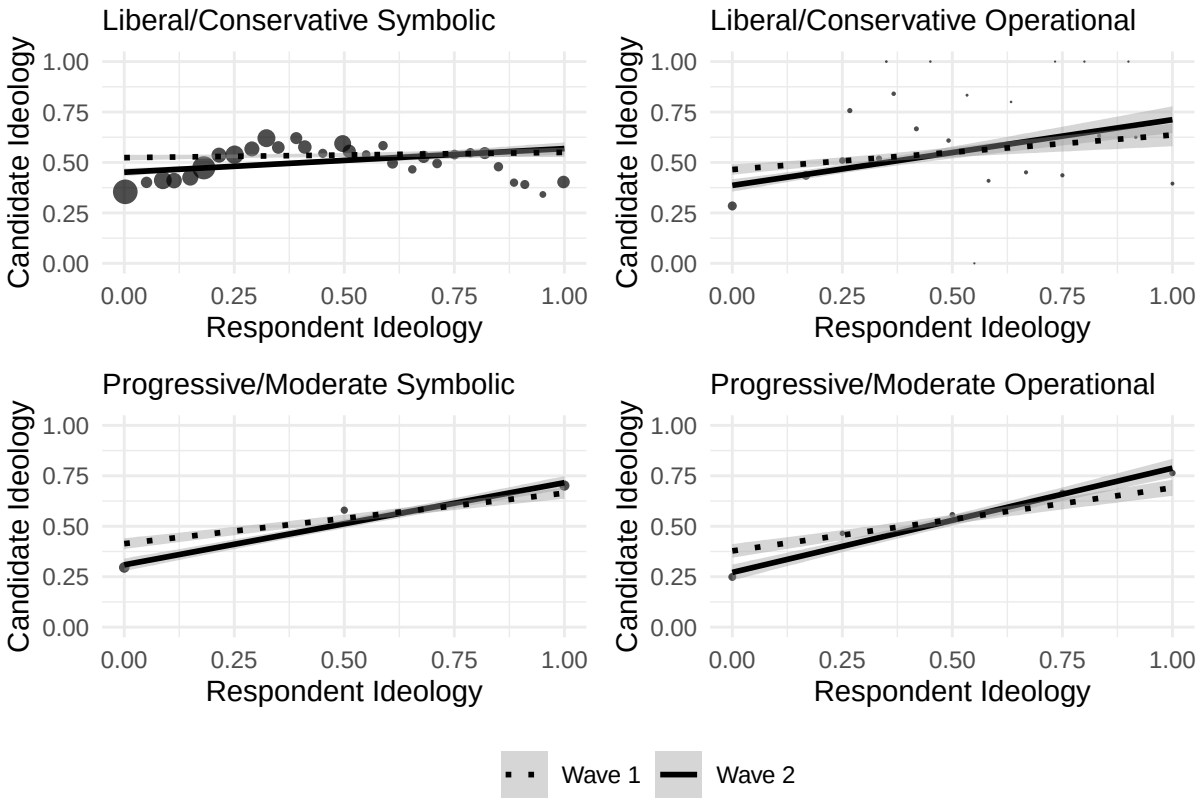

Figure 1: Respondent Ideology Regressed on the Ideology of Respondent's Vote Choice

Figure 2 allows one to see where the movement towards the aligned candidate is coming from. The three graphs in the figure show the relationships for three different groups of respondents: respondents who did not indicate a vote intent in wave 1, but did in wave 2 (Wave 1 Undecided), respondents whose vote intent in wave 1 is a candidate with a different ideology than their vote choice in wave 2 (Switchers), and respondents whose vote intent in wave 1 is the same ideology as their vote choice in wave 2 (Stayers). Since these groups condition on the dependent variable, Figure 2 should not be interpreted causally. Rather, it provides descriptive information that can inform inference. All three graphs use Liberal/Conservative Symbolic Ideology; the patterns are the same, but more exaggerated, for the other measures of ideology. See the Appendix for these additional graphs.

Respondents who do not switch their vote choice between waves, tend to already plan to vote for an ideologically-aligned candidate in wave 1. This pattern may help explain the positive relationship between respondent ideology and wave 1 vote intent ideology in Figure 1. Respondents who switch their vote choice between waves, tend to prefer ideologically-unaligned candidates in wave 1 and switch to ideologically-aligned candidates in wave 2. This pattern is reassuring, as it is consistent with learning candidate ideologies throughout the campaign and voting based on them. For respondents who were undecided in wave 1, I set their wave 1 vote choice ideology to the midpoint of the scale (0.5). By construction, this makes the wave 1 line a flat line at $y = 0.5$. I do this to keep these respondents in the analysis without letting an arbitrary left/right guess drive the slope: their true wave 1 preference is unobserved and could have gone either direction, so the neutral midpoint both retains them and limits their leverage in the regression analyses below. Notably, the relationship between respondent ideology and wave 2 vote choice ideology is positive, again showing evidence consistent with ideological voting. Since respondents who are undecided in wave 1 make up a much larger proportion of the overall respondent pool than switchers (42.9% for undecideds compared to 9.3% for switchers), the movement of the regression line in Figure 1 from wave 1 to wave 2 is likely driven mostly by individuals going from undecided to an ideologically-aligned vote.

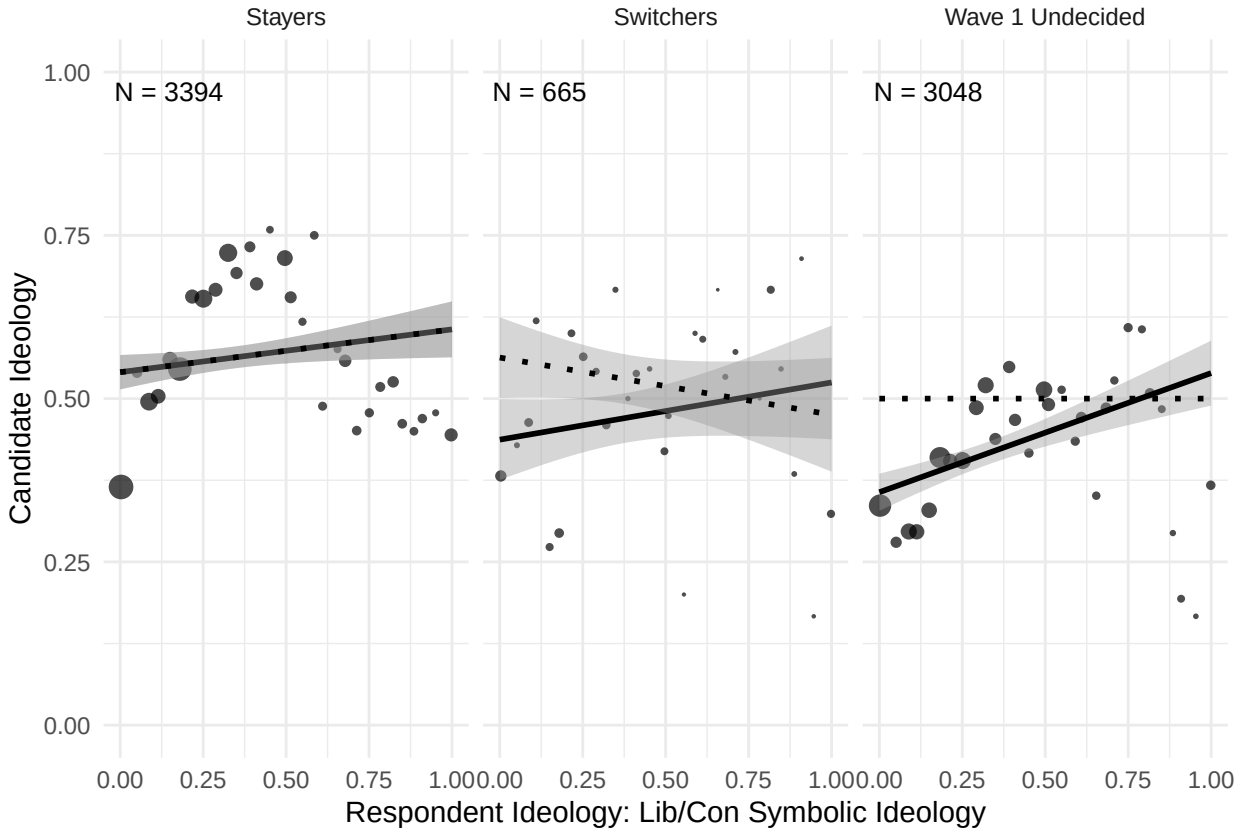

Figure 2: Respondent Ideology Regressed on the Ideology of Respondent's Vote Choice (Liberal/Conservative Operational Ideology)

The pattern of ideological alignment increasing from wave 1 vote intent to wave 2 vote choice can be succinctly captured by the following model:

$$y_{it} - y_{i(t-1)} = \alpha + \beta_1 x_{i(t-1)} + \epsilon_i$$

Again, y_{it} is the ideology of respondent i's vote choice on election day, $y_{i(t-1)}$ is the ideology of respondent i's vote intent in the pre-election wave, and $x_{i(t-1)}$ is respondent i's wave 1 ideology. Candidate ideology is 0 for relatively more liberal candidates, 1 for relatively more conservative candidates, and .5 if a respondent is undecided in wave 1. I include election fixed effects in the model to pick up on variation within primaries rather than across. I also run a lagged DV version of the model to make sure the coefficients are not being driven by regression to the mean.

$$y_{it} = \alpha + \beta_1 x_{i(t-1)} + \beta_2 y_{i(t-1)} + \epsilon_i$$

I follow Angrish and Pischke (2008) and treat the two models as marking the upper and lower bounds of the range within which the true effect falls.

For both models, β_1 reflects the likelihood that a liberal respondent switches their vote choice from undecided or a relatively more conservative candidate to a relatively more liberal candidate compared to the likelihood that a conservative respondent switches their vote choice from undecided or a relatively more conservative candidate to a relatively more liberal candidate. In other words, if β_1 is positive, on average, a respondent who is ideologically-aligned with a candidate is more likely to switch their vote choice to that candidate come wave 2 than a respondent who is ideologically-unaligned with that candidate. Table 2 reports β_1 for the lagged DV models and Table 3 reports β_1 for the difference models for each operationalization of ideology. I find consistent positive and significant coefficients across models. The magnitude of the coefficients in Table 2 vary, with Liberal/Conservative Symbolic Ideology producing the smallest coefficients and Liberal/Conservative Operational ideology producing the largest. On the low end, I find that respondents whose ideologies align with a candidate are 4-10% more likely to switch their vote to that candidate come wave 2 compared to respondents shows ideologies do not align with that candidate; On the high end, I find that aligned respondents are 25.1-38.8% more likely to switch.

Table 2: The relationship between respondent ideology and the ideology of second wave vote choice, controlling for the ideology of first wave vote intent.

	Lib/Con Symbolic Ideology	Lib/Con Operational Ideology	Prog/Mod Symbolic Ideology	Prog/Mod Operational Ideology
	DV: Ideology of Wave 2 Vote Choice			
	(1)	(2)	(3)	(4)
Respondent Ideology	0.112***	0.375***	0.206***	0.296***
	(0.025)	(0.044)	(0.018)	(0.027)
Ideology of Wave 1 Vote Intent	0.520***	0.593***	0.532***	0.571***
	(0.013)	(0.021)	(0.023)	(0.029)
Num.Obs.	7107	2111	1822	1179
R2	0.410	0.519	0.547	0.536
R2 Adj.	0.407	0.516	0.543	0.531
AIC	6646.6	1546.3	1235.7	837.5
BIC	6942.0	1642.4	1323.8	918.7

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: The relationship between respondent ideology and the difference between the ideology of second wave vote choice and first wave vote intent.

	Lib/Con Symbolic Ideology	Lib/Con Operational Ideology	Prog/Mod Symbolic Ideology	Prog/Mod Operational Ideology
	DV: Ideology of Wave 2 Vote Choice			
	(1)	(2)	(3)	(4)
Respondent Ideology	0.063*	0.252***	0.132***	0.191***
	(0.027)	(0.048)	(0.020)	(0.028)
Num.Obs.	7107	2111	1822	1179
R2	0.125	0.059	0.092	0.097
R2 Adj.	0.120	0.053	0.085	0.087

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Learning

One benefit of having panel data that enables within respondent analyses is that the models are not vulnerable to respondent-level confounders. However, we need to consider candidate-level confounders. The most likely candidate-level confounders are non-ideological factors, such as valence traits, that are more likely to be held by candidates with certain ideologies and appeal to respondents of that same ideology. For example, younger candidates in Democratic primaries may be more likely to hold more extreme ideologies and may be more appealing to respondents with more extreme candidates, not because they use age as a heuristic for ideology, but rather because they want someone "new" in office to shake things up. Though we control for some of these potential confounders in the models presented in the previous section, namely age, race, and prior government experience, there are many candidate characteristics we cannot control for. Given this, to increase confidence that the results are being driven by ideology I compare the voting behavior of respondents who learn candidates' ideologies between wave 1 and wave 2 and respondents who do not have knowledge of candidates' ideologies in wave 1 or wave 2. If ideology is motivating our results, we would expect to find significantly more ideological voting among respondents who learn candidates' ideologies.

In a study of Senate primary voters Hirano et al. (2015), use what they call a "minimal definition of knowledge" to test how informed their respondents are. They define "knowledge" as being able to place "at least one conservative candidate to the right of a liberal candidate" (102). I will refer to this as the "minimal knowledge test". If voters are able to place one relatively more conservative candidate to the right of one relatively more liberal candidate, they "pass" the "minimal knowledge test". Respondents are labeled as having learned candidate ideologies if they do not pass the minimal knowledge test in wave 1 but do pass in wave 2.

For this test, the ideology of primary candidates is determined using the same method as the voting models. Candidates are grouped into two categories: relatively more liberal candidates and relatively more conservative candidates. The former are coded as 0 and the latter are coded as 1. This is done at the primary election level, meaning that a relatively more liberal candidate in a Republican primary would likely be coded as a relatively more conservative candidate in a Democratic primary. I, along with research assistants, determined which candidates fit in which categories based on in-depth research we did on candidate positions in each primary. To measure voter perceptions of candidates, Hirano et al. (2015) use a typical 7-point ideology scale. Due to the possibility that primary voters may perceive all primary candidates to be at the extremes of a 7-point scale, while still perceiving ideological differences between the candidates, the "extremely liberal" to "extremely conservative" scales in the 2022 and 2024 surveys are presented to

respondents as slider questions that enable them to report small ideological differences between candidates. The slider question in the 2022 survey translates to a 70-point (1-7 in increments of .1) scale and in the 2024 survey translates to an 100-point scale (1-100 in increments of .1). Both are normalized between 0 and 1. If a respondent chose "unsure" or skipped this question, they are coded as not knowing the relative ideological positions of the candidates.

To determine whether respondents who learn candidate ideologies are more likely to cast an ideologically-aligned vote than respondents who do not know candidates' ideologies, I adjust the models used in the previous section to include an interaction term between respondent ideology and an indicator for learning.

$$y_{it} - y_{i(t-1)} = \alpha + \beta_1 x_{i(t-1)} + \beta_2 x_{i(t-1)} * l_i + \epsilon_i$$

$$y_{it} = \alpha + \beta_1 x_{i(t-1)} + \beta_2 x_{i(t-1)} * l_i + \beta_3 y_{i(t-1)} + \beta_4 y_{i(t-1)} * l_i + \epsilon_i$$

l_i is coded as 1 if the respondent does not pass the minimal knowledge test in wave 1 and passes in wave 2 ("learners") and is coded as 0 if the respondent does not pass the minimal knowledge test in wave 1 or wave 2 ("never knowers"). I run both the lagged and difference models, and include election fixed effects. β_1 measures the amount of ideological voting among respondents who do not know candidates' ideologies in either wave. β_2 measures the differential effect of ideology for learners, and thus, $\beta_1 + \beta_2$ indicates the amount of ideological voting among learners. Table 4 and Table 5 report the amount of ideological voting among learners ($\beta_1 + \beta_2$) and never knowers (β_1) for both the lagged and difference models, respectively.

Table 4: Comparing the relationship between respondent ideology and the ideology of second wave vote choice, controlling for first wave vote intent, of respondents who learned candidate ideologies and respondents who never knew candidate ideologies.

	Lib/Con Symbolic Ideology	Lib/Con Operational Ideology	Prog/Mod Symbolic Ideology	Prog/Mod Operational Ideology
	DV: Ideology of Wave 2 Vote Choice			
	(1)	(2)	(3)	(4)
Respondent Ideology	−0.150***	−0.022	0.047	0.005
	(0.037)	(0.086)	(0.040)	(0.062)
Ideology of Wave 1 Vote Intent	0.489***	0.603***	0.476***	0.517***
	(0.025)	(0.054)	(0.053)	(0.067)
Respondent Ideology x Learned Candidate Ideologies	0.471***	0.745***	0.300***	0.492***
	(0.044)	(0.102)	(0.054)	(0.083)
Ideology of Wave 1 Vote Intent x Learned Candidate Ideologies	−0.100**	−0.144	−0.112	−0.076
	(0.035)	(0.075)	(0.073)	(0.096)
Num.Obs.	4741	961	877	567
R2	0.402	0.465	0.506	0.473
R2 Adj.	0.396	0.454	0.496	0.457

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Comparing the relationship between respondent ideology and the difference between the ideology of second wave vote choice and first wave vote intent of respondents who learned candidate ideologies and respondents who never knew candidate ideologies.

	Lib/Con Symbolic Ideology	Lib/Con Operational Ideology	Prog/Mod Symbolic Ideology	Prog/Mod Operational Ideology
	DV: Ideology of Wave 2 Vote Choice - Ideology of Wave 1 Vote Choice			
	(1)	(2)	(3)	(4)
Respondent Ideology	−0.079*	−0.013	0.070	0.036
	(0.040)	(0.092)	(0.044)	(0.068)
Respondent Ideology x Learned Candidate Ideologies	0.397***	0.760***	0.239***	0.441***
	(0.048)	(0.109)	(0.060)	(0.091)
Num.Obs.	4741	961	877	567
R2	0.175	0.182	0.197	0.202
R2 Adj.	0.168	0.168	0.183	0.180

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4 indicates strong learning effects across all four measures of ideology. Learners are around 30-75% more likely to switch their vote to the more ideologically-aligned candidate come election day. Voters in the sample who never know candidate ideologies are no more likely to switch their vote to the aligned or unaligned candidate, except for Lib/Con Symbolic ideology where never knowers are 9-17% more likely to switch their vote to an ideologically *unaligned* candidate.

Though this instills more confidence that the observed shift towards more aligned candidates come election day is, in large part, being driven by ideology and not a potential confounder, the large impact learners have on the overall results raises an additional important question: do learners comprise a large enough percent of the primary electorate to contend that ideological voting is meaningfully contributing to election results?

Figure 3 plots the proportion of primary voters in each election that pass the minimal knowledge test in wave 1 against the proportion of primary voters in each election that pass the minimal knowledge test in wave 2. Each dot represents a different primary election and the size of the dots represent the number of respondents in each district. The dotted line is the 45-degree line; all dots above this line indicate elections where a greater proportion of primary voters passed the minimal knowledge test in wave 2 compared to wave 1. In other words, dots above this line indicate elections where at least some learning of candidate ideologies occurred between wave 1 and wave 2.[15] When respondents from all elections are combined, in wave 1 30% of primary voters are able to accurately place at least two candidates relative to one another ideologically; in wave 2 54% of primary voters are able to do so, suggesting that 24% of the overall sample of primary voters learned candidate ideologies between waves. Keeping in mind that not all voters who know candidate ideologies in wave 2 will vote based on them, it is reasonable to believe that the behavior of 25-50% of an electorate will have a substantial impact on election outcomes. Futhermore, these findings underscore the importance of efforts to further inform House primary electorates about candidate ideologies.

[15] Recall that in 2022 wave 1 data was collected 30 days before the primary election and in 2024 wave 1 data was collected 50 days before the primary election. In both years, wave 2 data was collected starting the day after the primary election.

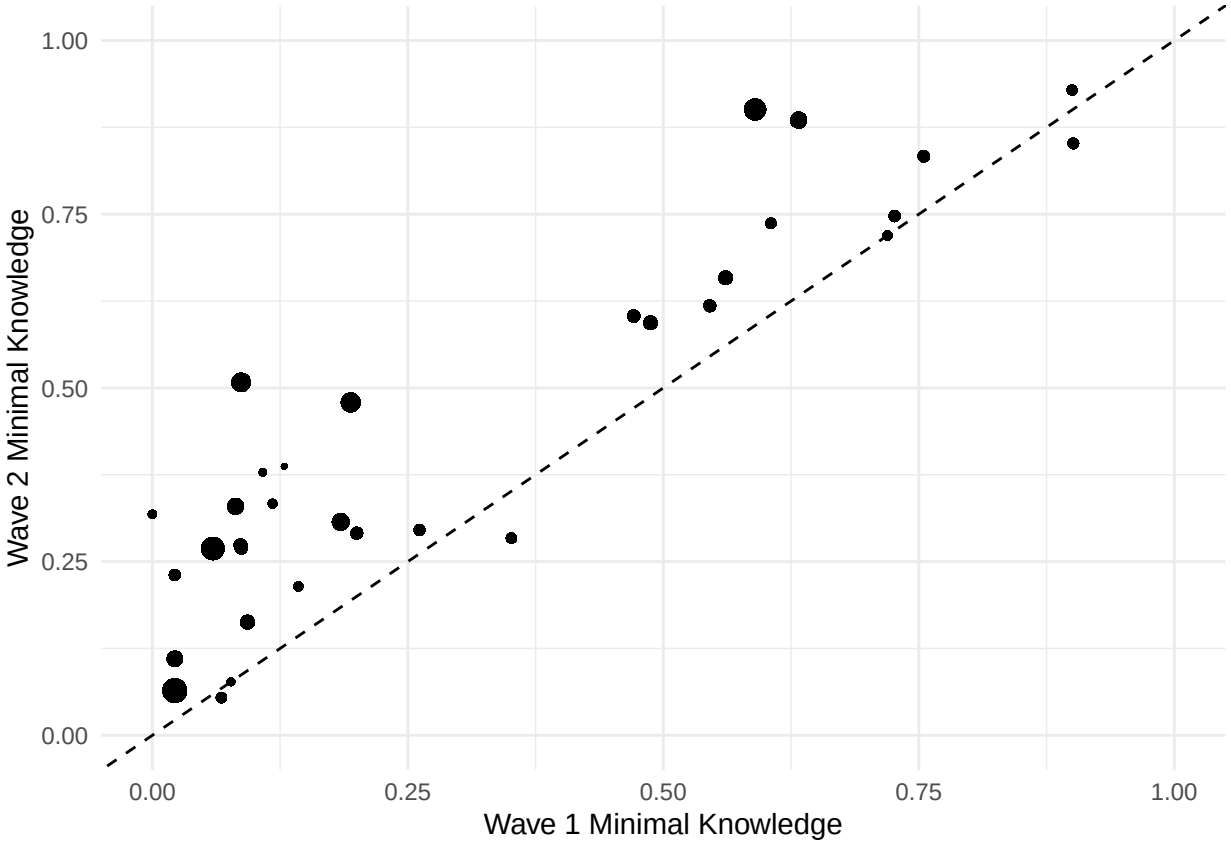

Figure 3: Minimal Knowledge Test

Impact of Parties

So far, I have presented evidence that ideological voting is a meaningful factor in House primary elections. In this section, I investigate whether ideological voting is meaningful in a specific subset of elections: elections where party organizations endorse a candidate. The party has many tools at its disposal to winnow the field in a way that benefits it's preferred candidate prior to the filing deadline, in a period of time referred to as the "invisible primary" (Aldrich 1987). This section considers the impact of election dynamics after the invisible primary where the party's two strongest tools are money and endorsements. I focus on endorsements because formal party PACs tend only to donate to candidates in House primaries under atypical circumstances, such as an incumbent facing a serious primary challenger. Which donations and independent expenditures to count as "party money" outside of formal party PACs is highly debatable. Conversely, party organization endorsements can be clearly defined and occur in around half of the elections in my sample (19 elections do not have a party endorsement and 21 elections do). The main impediment to the utility of party endorsements would be a lack of voter knowledge. However, voter knowledge of party endorsements is quite similar to knowledge of relative ideologies. 33% of voters in endorsed districts have knowledge of party endorsements in wave 1 and 58.4% of voters know who the party endorsed in wave 2. Thus, formal party endorsements present the cleanest path to investigate potential impacts of party involvement in House primaries after the invisible primary stage.

Formal party organizations do not endorse in every House primary. Thus, the prior results across all elections, with and without party endorsements, do not speak directly to whether ideological voting is meaningfully weaker when parties are actively involved. Simply re-estimating the individual-level models on subsets of elections with or without endorsements would be inappropriate, since endorsements are determined at the election level rather than the individual level. Instead, this section employs a series of election-level approaches to assess the relationship between party endorsements and ideological voting, and shows that overall levels of ideological voting are similar in endorsed and non-endorsed contests; party endorsements do not appear to weaken ideological voting.

First, I examine the distribution of ideological voting coefficients in elections with and without party endorsements. These estimates come from the baseline lagged dependent variable models used earlier in the paper:

$$y_{it} = \alpha + \beta_1 x_{i(t-1)} + \beta_2 y_{i(t-1)} + \epsilon_i$$

I run this model for each election individually. Figure 4 shows the overall distribution of these estimates, β_1, grouped by whether a party organization endorsed in the election. The dots size reflects the number of

respondents in each election.

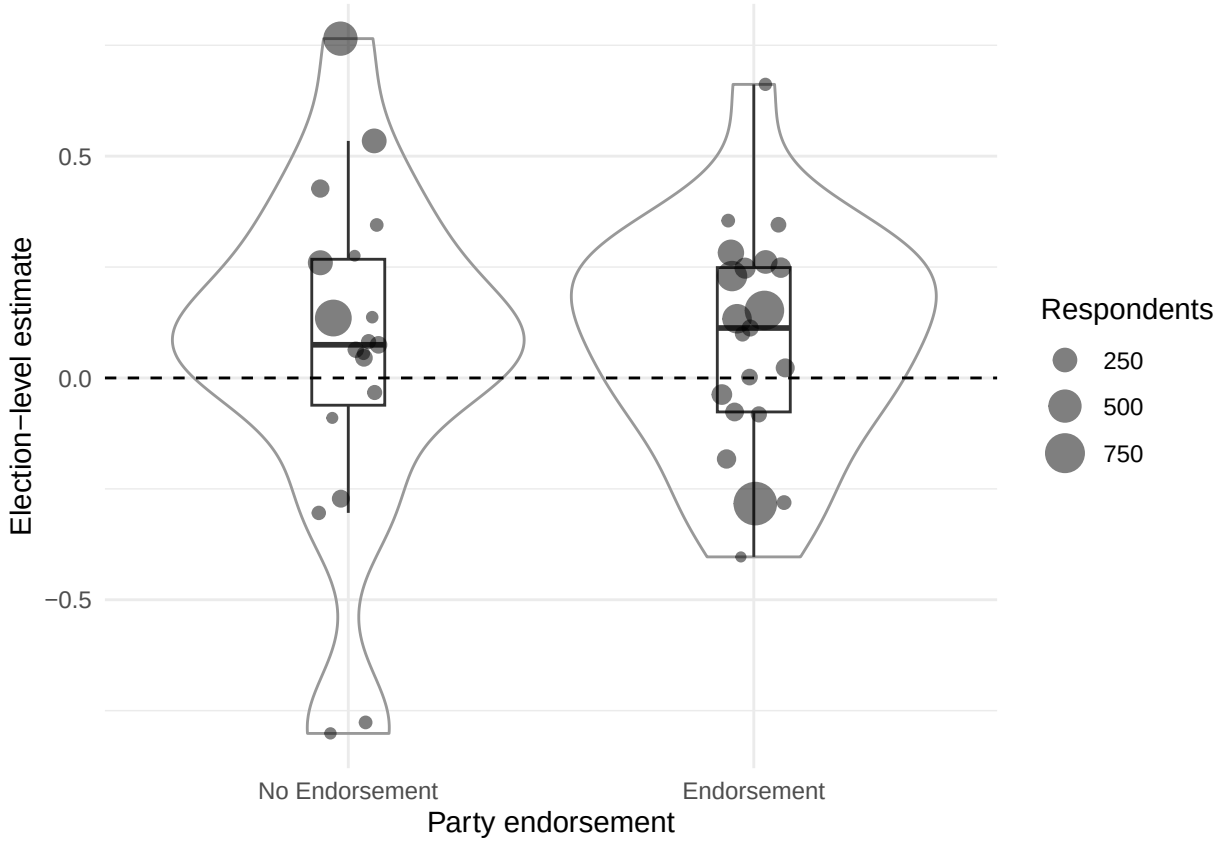

Figure 4: Distribution of ideological voting by party endorsement

The distributions in Figure 4 appear very comparable, suggesting little difference between ideological voting in elections with and without party endorsements. Table 6 formalizes this analysis further. It presents the results from a meta-analysis of each election's ideological voting estimate using a multilevel linear model:

$$y_{ij(t-1)} = \alpha + \beta_1 x_{ij(t-1)} + \beta_2 y_{ij(t)} + u_{0j} + u_{1j} x_{ij(t-1)} + \epsilon_{ij}$$

This model differs from model estimates reported in Figure 4 in that it estimates each parameter for individual i in election j, and introduces random intercepts for each election, u_{0j}, and random slopes for ideology in each election, $u_{1j} x_{ij(t-1)}$. Election-level random intercepts and random slopes for respondent ideology account for clustering of respondents within elections. This specification estimates (a) the average relationship between respondent ideology and wave 2 vote choice ideology, controlling for wave 1 vote choice, across all elections, and (b) the variation in this relationship across elections. Elections with more respondents contribute more information to the pooled estimate, while slopes from smaller elections are partially pooled toward the overall

Table 6: Multilevel linear model estimates of ideological voting by party endorsement

	DV: Vote Choice Ideology (Wave 2)	
	Lagged DV	Change Score
Respondent ideology (lib/con symbolic)	0.033	−0.001
	(0.077)	(0.071)
Party endorsed in Election	0.027	0.005
	(0.087)	(0.073)
Respondent Ideology × Party Endorsed	0.060	0.053
	(0.105)	(0.097)
Vote choice ideology (Wave 1)	0.513***	
	(0.021)	
Num.Obs.	7107	7107
R2 Marg.	0.168	0.001
R2 Cond.	0.368	0.143
ICC	0.2	0.1
Num.Groups (Elections)	40.000	40.000

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

average.

Table 6 presents the results of this model with an indicator for whether a party organization endorse a candidate in the election as an interaction term. The coefficient on Respondent Ideology X Party Endorsement indicates how much more/less ideological voting there is among elections where party's endorsed a candidate compared to elections where they did not endorse a candidate. The coefficient estimate is not significant, suggesting that in my sample there is not a difference in ideological voting between elections where party's do and do not endorse. One may reasonably argue that with only 40 elections total the model is under-powered. However, the coefficient estimate is positive, meaning that if it were statistically significant it would indicate *higher* levels of ideological voting *when parties endorse*. Consequently, the conclusion that, for the elections in my sample, there is not less ideologically voting when parties endorse is not merely a function of low power.

Table 7: The impact of learning candidate ideologies on ideological voting in primaries with and without party endorsements

	Lagged DV Models		Differnce Models	
	Party Endorse	No Party Endorse	Party Endorse	No Party Endorse
	(1)	(2)	(3)	(4)
Respondent Ideology	−0.139**	−0.146*	−0.056	−0.101
	(0.048)	(0.061)	(0.052)	(0.065)
Ideology of Wave 1 Vote Intent	0.461***	0.547***		
	(0.030)	(0.045)		
Respondent Ideology x Learned Candidate Ideologies	0.385***	0.537***	0.323***	0.462***
	(0.064)	(0.066)	(0.070)	(0.072)
Ideology of Wave 1 Vote Intent x Learned Candidate Ideologies	−0.072	−0.160*		
	(0.043)	(0.062)		
Num.Obs.	3097	1644	3097	1644
R2	0.411	0.373	0.188	0.152
R2 Adj.	0.407	0.364	0.182	0.141

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note:

All models include election fixed effects.

Table 8: The impact of learning party endorsements on ideological voting

	Lagged DV Model	Differnce Model
	(1)	(2)
Respondent Ideology	0.140*	0.082
	(0.062)	(0.067)
Ideology of Wave 1 Vote Intent	0.519***	
	(0.042)	
Respondent Ideology x Learned Party Endorsement	0.012	0.026
	(0.070)	(0.075)
Ideology of Wave 1 Vote Intent x Learned Party Endorsement	−0.092	
	(0.059)	
Num.Obs.	1782	1782
R2	0.432	0.256
R2 Adj.	0.426	0.248

* p < 0.05, ** p < 0.01, *** p < 0.001

Note:
All models include election fixed effects.

I find similarly null results when I examine the impact of learning on ideological voting and knowledge of party endorsements on ideological voting. Ideology is certainly not the only candidate characteristic that factors into voting decisions. It is possible that voters see party endorsements as a strong signal of candidate quality and are knowingly sacrificing optimal ideological-alignment for candidate quality. Table 7 and Table 8 begin to call this theory into question. Table 7 indicates no statistically significant difference in the effect of learning candidate ideologies on ideological voting between endorsed and non-endorsed elections. If voters in endorsed races were knowingly trading off ideological alignment for candidate quality, we would expect learning to have a weaker effect in those contexts. The absence of such a moderation effect suggests that increased attention to quality is unlikely to be driving the overall lower levels of ideological voting in elections with party endorsements. This reasoning is tested more directly, though on a smaller sample, in Table 8, which examines the impact of learning party endorsements on ideological voting. Once again, if voters in endorsed races were prioritizing candidate quality over ideology, we would expect those who learn party endorsements to engage in less ideological voting than those who do not. However, no such difference emerges.

Lastly, descriptive analyses of voter knowledge reveal knowledge levels consistent with finding no differences in ideological voting in elections with and without party endorsements. One may think that the presence of party endorsements reduces voter motivation to learn about non-endorsed candidates' ideologies. A party endorsement may be interpreted by voters as a signal that the endorsed candidate's quality is so superior to that of their opponents that further effort to learn about ideological differences is not worthwhile, as any such differences would be outweighed by the perceived quality gap. Therefore, the proportion of learners may be lower in endorsed races. When the data are weighted so each election contributes equally to the analysis, there is no statistically significant difference in wave 1 knowledge, wave 2 knowledge, or net learning by party endorsement (see Figure 5).[16]

[16]Net learning is defined as the share of respondents who do not pass the minimal knowledge test in wave 1 but do in wave 2, minus the share who pass in wave 1 but not in wave 2 (learners minus forgetters).

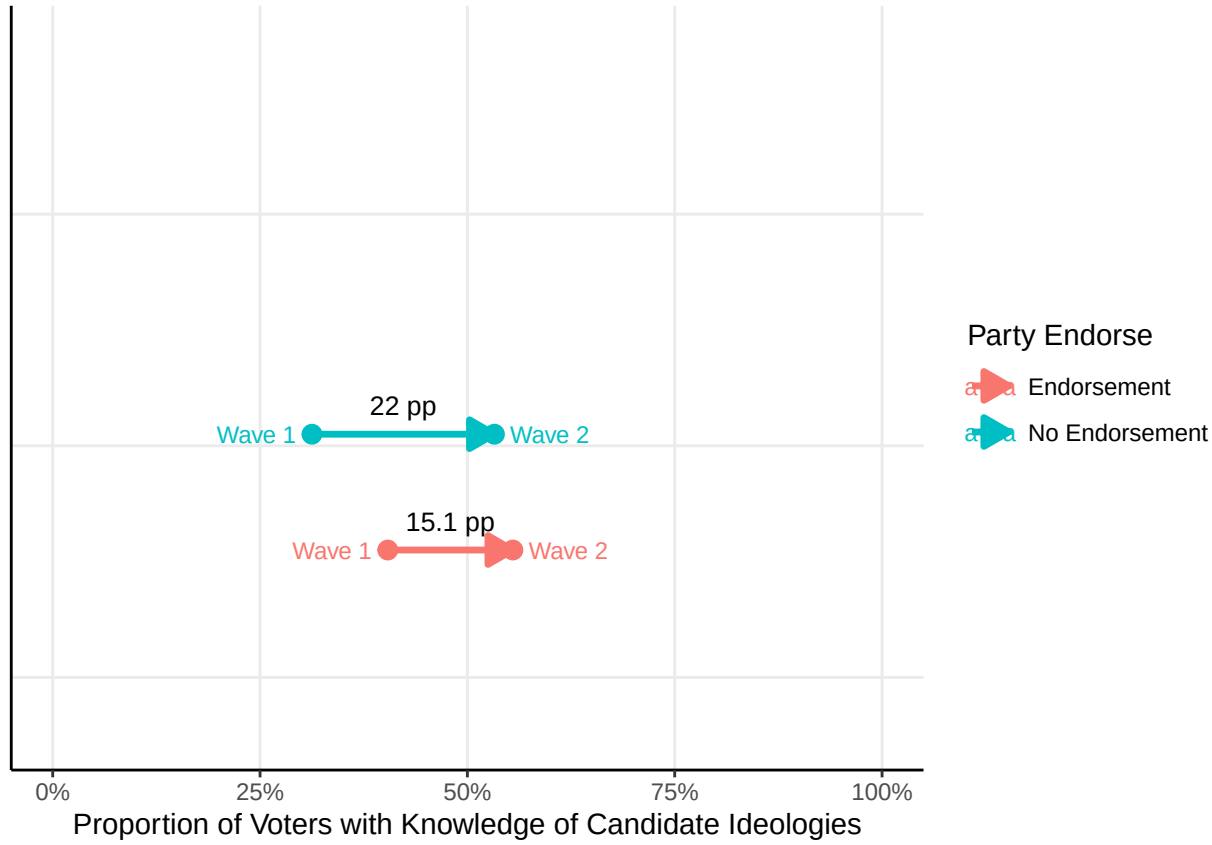

Figure 5: Respondent knowledge of candidate ideologies in elections with and without party endorsements

I also examine differences in three election-level characteristics that are correlated with respondent knowledge of candidate ideologies, namely average campaign spending per candidate, the proportion of candidates in the election with prior office-holding experience, and the number of candidates running in the election, and find limited divergence of elections where party organizations do and do not endorse. Figure 6 presents evidence of intuitive relationships between these election-level characteristics and knowledge. There is a positive relationship between average per candidate campaign spending and wave 2 knowledge (more spending leads to more information availability leads to more knowledge), a positive relationship between the proportion of candidates in a given primary with prior office-holding experience and wave 2 knowledge (holding prior office increases the likelihood that voters are already familiar with the candidate which increases the likelihood of ideological knowledge), and a negative relationship between the number of candidates in a race and wave 2 knowledge (more candidates increases the cognitive load on voters and makes it more difficult to retain knowledge of candidate ideologies). Table 9 compares each of these election-level characteristics with and without party endorsements at the election-level. There is no statistically significant difference in campaign spending or candidates with prior office-holding experience between elections with and without party endorsements. There is a statistically significant difference between number of candidates, but elections *without* party endorsements have *more* candidates on average. However, this would suggest less knowledge of candidate ideologies in elections without party endorsements, not more. Thus, direct examinations of ideological learning, as well as examinations of possible mechanisms that could lead to differences in ideological voting, are all consistent with similar levels of ideological voting across elections with and without party endorsements. It appears that this party tool does not mitigate ideological voting.

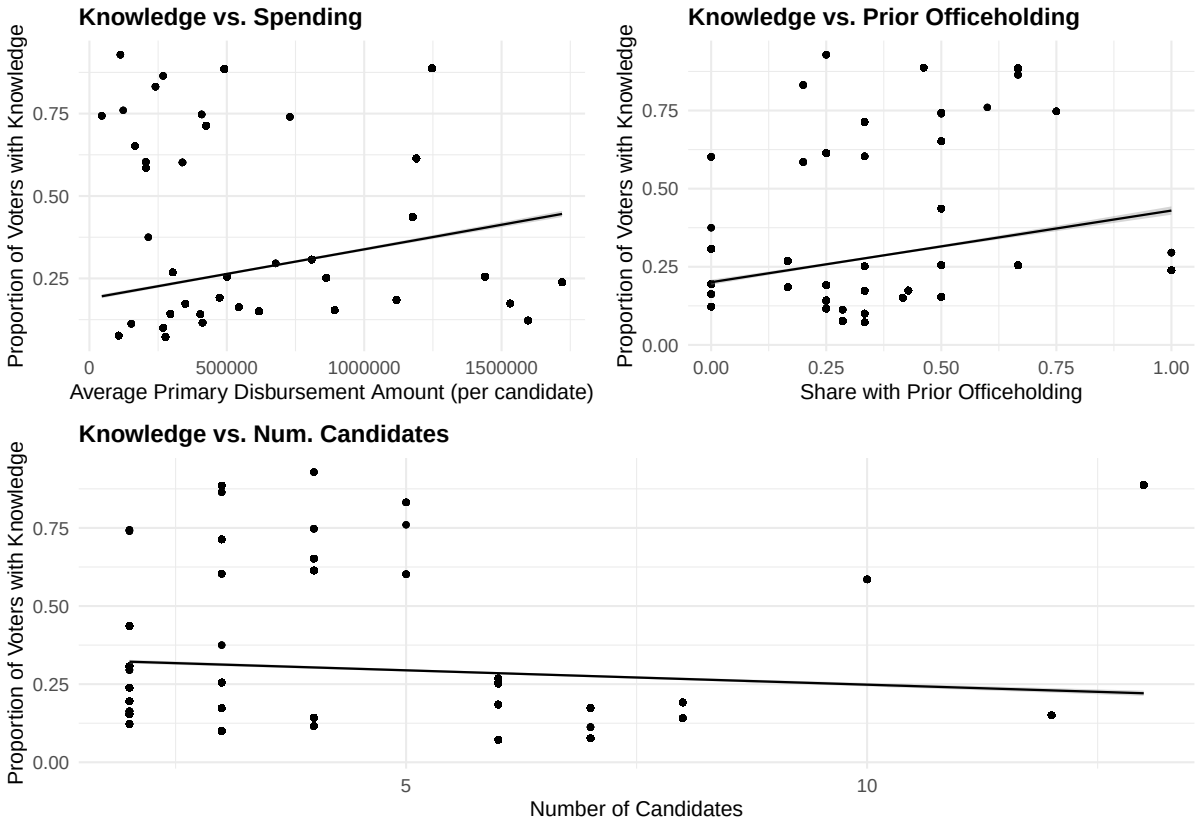

Figure 6: Relationship between wave 2 respondent knowledge of candidate ideologies and election-level characteristics

Table 9: Means of election-level characteristics associated with voter knowledge

	No Endorsement		Endorsement		
Variable	Mean	N	Mean	N	Diff in Means (Welch SE)
Proportion of Cands with Prior Office	0.307	19	0.410	21	0.103 (0.077)
Number of Candidates	5.632	19	3.714	21	-1.917* (0.843)
Average Spending Per Candidate	750733.913	19	648248.771	21	-102485.142 (168847.736)

$p < .001$ ***, $p < .01$ **, $p < .05$ *

Discussion

There are two main ways congressional primaries can contribute to polarization: through parties and through voters. Though the literature has largely ruled out voter-driven polarization, few studies directly examine voting behavior in congressional primaries. More recently, motivated by the UCLA school theory of parties (Bawn et al. 2012), the idea of elite-driven primary polarization has gained traction. This theory argues that parties mainly care about policy and use their influence in primary elections to nominate the candidate that holds the party's preferred policy positions. Under this theory, congressional primary voters end up voting for the party-preferred candidate, regardless of the voters' own preferences, due to the difficulty of ideologically differentiating intra-party candidates and the name recognition parties create for their preferred candidate through various avenues of promotion. Using a first-of-its-kind dataset of ~8,000 House primary voters in 40 primary elections across 10 states and 2 primary cycles, I find evidence of ideological voting and voter autonomy, even when party organizations are directly involve in election dynamics through endorsements.

The presence of ideological voting in House primaries presents an interesting tension. On one hand, it softens the idea that direct primaries in the United States put nominations to the voters in name-only, whereas in practice parties are still able to pull the strings necessary to elect their preferred candidates. The impact of the invisible primary on election outcomes should not be understated, but, the bluntest devise parties have to sway election outcomes after the invisible primary, endorsements, does not seem to lessen ideological voting, even with over half of voters in the sample having knowledge of party endorsements. In this sense, the Progressive Era dream of "the people" having more electoral control appears alive and well. But, on the other hand, only a limited portion of the electorate chooses to exercise this control in primary elections. Though evidence on whether this subset of voters actually is more ideologically extreme than the party

median is mixed, a strong presence of ideological voting provides motivation to continue to investigate this in the context of whether House primary voters are contributing to polarization. The more ideological control voters have over candidates nominated to the general election, the more likely it is that primary voters are a driver of polarization in the United States (conditional on the extremity of the primary electorate).

Future research should examine whether parties or voters tend to prefer more extreme candidates. If voters prefer the more extreme candidate, then party influence potentially serves to mitigate the polarizing effect of primary voters. If parties prefer the more extreme candidate, then increasing voter control further with potential reforms to the invisible primary stage may mitigate the potential polarizing effect of primaries. It is also important for future research to investigate the reasons why party organizations get involved in some primary elections but not others. Why do some state parties, like the California Democrats, endorse in almost every primary, while other state parties endorse in none? Understanding whether this is a strategic choice, if so, what the underlying strategy is, and if not, the reason for the choice, will help determine the potential implications of these decisions. Additionally, future research should examine the process through which party organizations choose who to endorse. Whose preference does the endorsement actually represent?

This paper underscores the importance of studying primaries at the voter-level. More survey data on House primary voters needs to be collected in order to see whether these results remain when additional primaries are added to the dataset. Because of the lack of survey data on House primary voters we do not know whether the dynamics described in this paper are unique to the 2022 and 2024 primary cycles. Since we cannot go back in time, the best we can do is continue this type of data collection in the future so the role of House primary voters does not go unexamined.

References

Achen, C. H., & Bartels, L. M. (2016). Democracy for realists. Princeton University Press.

Ahler, D. J., Citrin, J., & Lenz, G. S. (2016). Do Open Primaries Improve Representation? An Experimental Test of California's 2012 Top-Two Primary. Legislative Studies Quarterly, 41(2), 237-268.

Allcott, H., Gentzkow, M., Mason, W., Wilkins, A., Barberá, P., Brown, T., & Tucker, J. A. (2024). The effects of Facebook and Instagram on the 2020 election: A deactivation experiment. Proceedings of the National Academy of Sciences, 121(21), e2321584121.

Aragones, E., & Neeman, Z. (2000). Strategic ambiguity in electoral competition. Journal of Theoretical Politics, 12(2), 183-204.

Bawn, K., Cohen, M., Karol, D., Masket, S., Noel, H., & Zaller, J. (2012). A theory of political parties: Groups, policy demands and nominations in American politics. Perspectives on Politics, 10(3), 571-597.

Bawn, K., DeMora, S. L., Dowdle, A., Hall, S., Myers, M. E., Patterson, S., & Zaller, J. (2019). Policy voting in US House primaries. Journal of Elections, Public Opinion and Parties, 29(4), 533-549.

Binder, S. A. (1999). The dynamics of legislative gridlock, 1947–96. American Political Science Review, 93(3), 519-533.

Bullock, W., & Clinton, J. D. (2011). More a molehill than a mountain: the effects of the blanket primary on elected officials' behavior from California. The Journal of Politics, 73(3), 915-930.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1980). The american voter. University of Chicago Press.

Chen, K. P., & Yang, S. Z. (2002). Strategic voting in open primaries. Public Choice, 112(1), 1-30.

Cohen, M., Karol, D., Noel, H., & Zaller, J. (2008). The party decides: Presidential nominations before and after reform. University of Chicago Press.

Crowder-Meyer, M. (2013). Gendered recruitment without trying: How local party recruiters affect women's representation. Politics & Gender, 9(4), 390-413.

Desmarais, B. A., La Raja, R. J., & Kowal, M. S. (2015). The fates of challengers in US house elections: The role of extended party networks in supporting candidates and shaping electoral outcomes. American Journal of Political Science, 59(1), 194-211.

Downs, A. (1957). An economic theory of political action in a democracy. Journal of political economy, 65(2), 135-150.

Free, Lloyd A., and Hadley Cantril. 1967. The Political Beliefs of Americans. New Brunswick, NJ: Rutgers University Press

Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. Nature human behaviour, 4(5), 472-480.

Grofman, B., & Norrander, B. (1990). Efficient use of reference group cues in a single dimension. Public Choice, 64(3), 213-227.

Hacker, J. S. (2004). Privatizing risk without privatizing the welfare state: The hidden politics of social policy retrenchment in the United States. American Political Science Review, 98(2), 243-260.

Hacker, J., & Pierson, P. (2006, June). Still off topic: a reply to Pitney's rejoinder. In The Forum (Vol. 4, No. 1, p. 000010220215408884119). De Gruyter.

Hayes, D., & Lawless, J. L. (2021). News hole: The demise of local journalism and political engagement. Cambridge University Press.

Hayes, D., & Lawless, J. L. (2021). News hole: The demise of local journalism and political engagement. Cambridge University Press.

Hassell, H. J. (2016). Party control of party primaries: Party influence in nominations for the US senate. The Journal of Politics, 78(1), 75-87.

Hassell, H. J. (2023). Party elite engagement and coordination in House primary elections: a test of theories of parties. American Journal of Political Science, 67(2), 307-323.

Henderson, J. A. (2013). Downs' revenge: elections, responsibility and the rise of congressional polarization. University of California, Berkeley.

Hill, S. J. (2015). Institution of nomination and the policy ideology of primary electorates. Quarterly Journal of Political Science, 10(4), 461-487.

Hirano, S., Lenz, G. S., Pinkovskiy, M., & Snyder Jr, J. M. (2015). Voter learning in state primary elections. American Journal of Political Science, 59(1), 91-108.

Hirano, S., & Snyder Jr, J. M. (2019). Primary elections in the United States. Cambridge University Press.

Hirano, S., Snyder Jr, J. M., Ansolabehere, S. D., & Hansen, J. M. (2010). Primary elections and partisan polarization in the US Congress. Quarterly Journal of Political Science.

Jefferson, H. (2020). The curious case of Black conservatives: construct validity and the 7-point liberal-conservative scale. Available at SSRN 3602209.

Koger, G., Masket, S., & Noel, H. (2009). Partisan webs: Information exchange and party networks. British Journal of Political Science, 39(3), 633-653.

Lee, F. E. (2009). Insecure majorities: Congress and the perpetual campaign. University of Chicago Press.

Lelkes, Y., Sood, G., & Iyengar, S. (2017). The hostile audience: The effect of access to broadband internet on partisan affect. American Journal of Political Science, 61(1), 5-20.

Lenz, G. S. (2012). Follow the leader?: how voters respond to politicians' policies and performance. University of Chicago Press.

Levendusky, M. (2009). The partisan sort: How liberals became Democrats and conservatives became Republicans. University of Chicago Press.

Meirowitz, A. (2005). Keeping the other candidate guessing: Electoral competition when preferences are private information. Public Choice, 122(3), 299-318.

Levendusky, M. (2009). The partisan sort. University of Chicago Press.

Lupia, A. (1994). Shortcuts versus encyclopedias: Information and voting behavior in California insurance reform elections. American Political Science Review, 88, 63-76.

Mann, T. E., & Ornstein, N. J. (2012). Let's just say it: The Republicans are the problem. Washington Post, 27.

Mason, L. (2015). "I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization. American journal of political science, 59(1), 128-145.

McCarty, N. (2007). The policy effects of political polarization. The transformation of American politics: Activist government and the rise of conservatism, 223-55.

McCarty, N., Poole, K. T., & Rosenthal, H. (2006). Polarized America: The dance of ideology and unequal riches. mit Press.

McGhee, E., Masket, S., Shor, B., Rogers, S., & McCarty, N. (2014). A primary cause of partisanship? Nomination systems and legislator ideology. American Journal of Political Science, 58(2), 337-351.

Norrander, B. (1989). Ideological representativeness of presidential primary voters. American Journal of Political Science, 570-587.

Oak, M. P. (2006). On the role of the primary system in candidate selection. Economics & Politics, 18(2), 169-190.

Page, B. I. (1976). The theory of political ambiguity. American Political Science Review, 70(3), 742-752.

Poole, K. T., & Rosenthal, H. (2000). Congress: A political-economic history of roll call voting. Oxford University Press, USA.

Popp, E., & Rudolph, T. J. (2011). A tale of two ideologies: Explaining public support for economic interventions. The Journal of Politics, 73(3), 808-820.

Porter, R., Treul, S. A., & McDonald, M. (2024). Changing the Dialogue: Descriptive Candidacies and Position Taking in Campaigns for the US House of Representatives. The Journal of Politics, 86(2), 458-474.

Shepsle, K. A. (1972). The strategy of ambiguity: Uncertainty and electoral competition. American Political science review, 66(2), 555-568.

Stimson, James A. 2004. Tides of Consent: How Public Opinion Shapes American Politics. New York: Cambridge University Press.

Woon, J. (2018). Primaries and candidate polarization: behavioral theory and experimental evidence. American Political Science Review, 112(4), 826-843.