

Do Primaries Work?
Bayesian Causal Models of Partisan Ideology and Congressional
Nominations

By
Michael G. DeCrescenzo

A dissertation submitted in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY (POLITICAL SCIENCE)
at the
UNIVERSITY OF WISCONSIN–MADISON,
2020

Approved by the thesis committee on the oral defense date, **TBD**

Barry C. Burden (Chair), Professor of Political Science

Kenneth R. Mayer, Professor of Political Science

Eleanor Neff Powell, Associate Professor of Political Science

Alexander M. Tahk, Associate Professor of Political Science

Michael W. Wagner, Professor of Journalism and Mass Communications

Abstract

In contemporary electoral politics in the U.S., primary elections are widely believed to play a crucial role. Many scholars believe that primary election competition is the standout reason why classic predictions from formal models of electoral competition—that candidates take ideological positions near the median voter—fail to manifest in the real world. The general election context provides incentives for candidates to take centrist policy positions, but candidates must win their party's nomination before advancing to the general election. Because primary elections take place predominantly among voters of one political party affiliation, and because those voters tend to hold strongly partisan beliefs about political issues, candidates feel more acute incentives to take strong partisan stances on issues rather than moderate stances even amid stiff general election competition.

This story of primary elections and representation is widely believed, but is it true? Despite its prominence, the empirical evidence is unclear. The theory rests on a notion that voters make informed choices in primary elections by consulting their policy preferences and choosing the candidate with the closest policy platform. Past research has been unable to operationalize key constructs in this prediction, or it has operationalized the wrong constructs. Candidates should take more extreme positions when the primary constituency has a stronger preference for ideologically extreme policy, but studies have not directly measured the policy preferences of partisans within a candidate's district. Further, districts where partisans hold more extreme preferences should nominate candidates with more extreme campaign positions as well, but methods for estimating candidates' ideological positions have been incompletely applied to the study of primaries. Moreover, because primary elections are characterized by low levels of voter information and the partisanship of candidates is held largely constant, non-policy forces such as candidate valence and campaign spending may be more powerful than in general elections. For these reasons, the proposition that primary

elections advance the ideological interest of local partisan voters is theoretically contestable.

This dissertation develops and applies new Bayesian approaches for estimating both constructs that have yet eluded the study of primary politics: the preferences of partisan voters as a group and the campaign positioning of primary candidates. With these estimates in hand, I explore the relationship between local partisan preferences and primary candidate positions. Do primary candidates position themselves relative to partisan primary voters, and is the relative extremism of partisan constituencies related to the ideological positions of the candidates they nominate?

Contents

Abstract	i
Contents	iii
List of Tables	v
List of Figures	vi
Acknowledgments	viii
1 Introduction: Policy Ideology and Congressional Primaries	1
1.1 Policy Preferences and the Strategic Positioning Dilemma	3
1.2 Does the Strategic Positioning Dilemma Describe Primary Representation? .	12
1.3 Project Outline and Contributions	27
2 Hierarchical IRT Model for District-Party Ideology	35
2.1 Testing the Model with Simulated Data	35
2.2 Data Sources	35
2.3 Model Results	35
3 Bayesian Causal Inference	37
4 How District-Party Ideology Affects Primary Candidate Positioning: A Bayesian De-Mediation Model	38
4.1 Candidate Positioning and Voters' Policy Preferences	40
4.2 The Causal Effect of District-Party Ideology	54

4.3 Findings	73
4.4 Discussion	84
5 District-Party Ideology and Primary Election Outcomes	86
References	87

List of Tables

4.1	Sample sizes in all estimated models.	74
-----	---	----

List of Figures

1.1	Non-identifiability of partisan group preferences from district vote shares.	22
1.2	The relationship between average ideological self-placement and district vote share in congressional districts.	26
2.1	Posteriors	36
4.1	Topline relationship between district-party ideology and candidate positions. . .	48
4.2	District-party ideology and candidate positioning for incumbents, challengers, and open-seat candidates.	50
4.3	District-party ideology and candidate positioning in 2012, 2014, and 2016.	51
4.4	Average CF scores in states with closed-semi-closed, and open primary rules. . .	52
4.5	District-party ideology and candidate positioning in states with different primary rules.	54
4.6	Key variables in the strategic positioning theory.	56
4.7	Causal diagram of the strategic positioning dilemma	57
4.8	Causal graphs describing the modeling problem and sequential g estimation. The stage 1 graph identifies the effect of the past district voting (V_d) on candidate positioning (CF_i). The stage 2 treatment-outcome model subtracts the district vote effect from candidate positions and identifies the effect of district-party ideology $\bar{\theta}_g$ on the demediated CF score $b(CF)_i$, which is equivalent to the controlled direct effect on the raw CF score. The final graph shows where unadjusted confounders violate the nonparametric causal identification assumptions in stage 1 ($U1$) and in stage 2 ($U2$).	61
4.9	Prior and posterior distributions for ideal points in sequential g model	75

4.10	Sequential g results for Democratic candidates.	76
4.11	Sequential g results for Republican candidates.	79
4.12	Sequential g results for incumbents, challengers, and open seat candidates. . . .	82
4.13	Sequential g results for candidates in states with closed, semi-closed, and open primary systems. Points and intervals are variational point estimates and 90 percent intervals from the approximate posterior distribution.	83

Acknowledgments

This project is better thanks to the advice and suggestions of my committee, Scott Straus and the students in the 2017 prospectus workshop, other Burden advisees (Levi Bankston, Jordan Hsu, Matt Shor, and Rochelle Snyder), David Canon, Devin Judge-Lord, Evan Morier, Blake Reynolds, and Marcy Shieh. Researchers at other institutions provided suggestions and moral support about the ideas in this project as well: Devin Caughey, David Doherty, Andrew Heiss, Seth Hill, Shiro Kuriwaku, TJ Mahr, Steve Miller, Jacob Montgomery, Arman Oganisian, John Poe, Rachel Porter, Jonathan Robinson, Sarah Treul, and Chris Warshaw. Extra special thanks to researchers whose freely-provided data made the project possible: Robert Boatright, Adam Bonica, and Ella Foster-Molina, Seth Masket, Nolan McCarty, Eric McGhee, Vincent Moscardelli, Steven Rogers, Boris Shor, Clifford Vickrey, and the teams who make the ANES and CCES surveys possible. I received financial support from the Elections Research Center at the University of Wisconsin–Madison, and technical support from UW’s Social Science Computing Cooperative, Garrick Aden-Buie, and Matthew Kay. Thanks also to Deb McFarlane for her help stewarding other students and myself through the bureaucracy of dissertating.

Thanks also to my parents, who kept so many pieces of my life together during this process. Thanks to Ricky O’Connor, who connected me to my first job after graduate school that gave me the push to finish the project and move on to the next thing. Thanks finally to some close friends who made my time in Madison better: Shaan Amin, Josh Cruz, Micah Dillard, José Luis Enríquez Chiñas, Caileigh Glenn, Ryan Hinchey, Jordan Hsu (again), Devin Judge-Lord (again), Amy Kawleski, Sara Linski, Richard Loeza, Anna Meier, Evan Morier (again), Erin Nelson, Anna Oltman, Madeline Vogt, and Erin Zwick.

1

Introduction: Policy Ideology and Congressional Primaries

Elections are the foremost venue for citizens to influence government actors and public policy. Classic theories of voting suggest that citizens weigh the policy positions of alternative candidates and vote for the candidate whose platform most closely aligns with their own preferences (Downs 1957). Political parties simplify the voter's calculations by providing a powerful heuristic in the form of the party label, enabling voters to infer candidates' values and issue positions without expending the effort to thoroughly appraise each campaign (Campbell et al. 1960; Green, Palmquist, and Schickler 2002; Rahn 1993).

The rise of partisan polarization, however, has complicated the role of parties in U.S. politics. Although citizens, journalists, pundits, and even elected leaders frequently bemoan the bitter rhetoric and legislative gridlock that has accompanied the widening partisan divide, political scientists have noted several positive consequences to polarization. Compared to the parties of the early- and mid-1900s that political scientists believed were too similar to provide voters with meaningful choices (American Political Science Association 1950), the Democratic and Republican Parties of recent decades have taken divergent and oppositional stances across a greater number of policy issues. As a result, voters can more easily differentiate the policy platforms of the two parties in order to vote consistently with their political values. Voters in turn became more thoroughly sorted into partisan groups that represent

distinct ideological viewpoints in American politics, holds beliefs across multiple issues that are more ideologically consistent, think more abstractly about the ideological underpinnings of issue stances, and participate more in politics than they did in the past (Abramowitz and Saunders 1998; Fiorina, Abrams, and Pope 2005a; Layman and Carsey 2002; Levendusky 2009).

Even as polarization has strengthened many aspects of political representation between the two parties, it may have troubling effects on representation within the two major parties. The typical voter is a partisan who intends to cast her ballot for her preferred party, whoever that candidate may be (Bartels 2000; Petrocik 2009). As party-line voting increases, voters are more thoroughly captured by their loyalties. A partisan voter's choices are locked in long before Election Day. Candidates from her preferred party have already been selected through a nomination process, and she may be more likely to abstain from voting when faced with an undesirable candidate than she is to vote for a different party (Hall and Thompson 2018). Recent research supports this notion of capture amid polarization—when voters must choose between polarized candidates, they become less responsive to candidates' actual platforms and instead are more influenced by motivated reasoning and partisan teamismanship (Rogowski 2016). Voters relax their substantive scrutiny of candidates to cast low-cost votes for their own party, weakening the influence of *policy* as a separate consideration from partisanship.

This presents an important problem for our understanding of how elections contribute to the representation of voter preferences in government. Elections are intended to be a voter's choice over alternative political values to be expressed in government, but if the choice of candidates does not present the average partisan voter with realistic alternatives, how should we think about the “representation” of these voters' actual policy preferences? If general elections provide an ever-coarsening choice over policy priorities, does the U.S. electoral system incorporate voters policy preferences in other ways?

When the choice before voters in the general election does not present realistic alterna-

tives, political scientists naturally shift their focus to the nomination of partisan candidates. V.O. Key, for example, studied Democratic Party dominance in the American South, asking if competition within the party could provide a quality of representation similar to two-party competition (Key 1949). Although scholars are right to examine within-party competition, focusing on contexts of single-party dominance is a serious limitation. Even in races between viable candidates from both major parties, within-party competition plays a crucial role simply due to the fact that partisan voters almost certainly cast a vote for their own party. Rank-and-file partisan constituents are all but captured. If they are to express their policy preferences through the act of voting, their voices may register as relatively weak because they present little electoral risk to their party in the general election. The nomination stage—the primary election in particular—remains an important venue for the representation of partisans’ policy views, whether the general election is closely contested or not.

1.1 Policy Preferences and the Strategic Positioning Dilemma

This dissertation is chiefly concerned with the policy preferences of partisan voters and their role in electoral representation through Congressional primary elections. The study of American electoral politics has not ignored the representational function of primary elections (Aldrich 2011; Cohen et al. 2009; Geer 1988; Norrander 1989; Sides et al. 2018), but as I discuss below, the quantifiable impact of primary voters’ policy preferences in government is a startlingly open question. Several existing studies have examined other aspects of representation through House primaries, such as the introduction of the direct primary (Ansolabehere et al. 2010), how candidates position themselves in response to the presence or threat of primary challenges (Brady, Han, and Pope 2007; Burden 2004; Hirano et al. 2010), and how primary nomination rules affect elite polarization (Hirano et al. 2010; McGhee et al. 2014; Rogowski and Langella 2015). Though these studies address interesting

aspects of electoral representation and party competition, they cannot speak directly to the influence of voter's policy preferences on (1) the positioning of House primary candidates and (2) the outcomes of House primary elections.

The absence of voter preferences from the empirical study of primaries is troubling because they play a crucial role in the dominant theory that relates representation to primary politics. Although the Downsian model of candidate positioning explains the incentives for candidates to stake out moderate policy positions to cater to the ideological “median voter” (Downs 1957), candidates behave differently in the real world. Instead, candidates engage in highly partisan behavior and take divergent issue stances even on salient local issues and in closely competitive districts (Ansolabehere, Snyder, and Stewart 2001; Fowler and Hall 2016). But why? Scholars and political observers have argued that because competing in the general election requires each candidate to clinch their party's nomination contest, these candidates face a combination of convergence-promoting and divergence-promoting incentives. Primary elections tend to be dominated by partisan voters who are more attentive to politics, hold more non-centrist issue preferences, and “weight” candidates' issue positions more heavily than the average voter in the general election.¹ As a result, the risk that a candidate is defeated in the primary for being too moderate may outweigh the risk of losing the general election for being too partisan. The conflicting incentives imposed by partisan constituency and the general election constituency creates a “strategic-positioning dilemma” that leads candidates to take divergent issue stances rather than targeting a district median voter (Aldrich 1983; Brady, Han, and Pope 2007; Burden 2001; Hill 2015).

The strategic positioning dilemma (SPD) is a central theoretical feature of this project,

¹Primary elections are not *entirely* partisan affairs. States vary in their regulations that primaries be “closed” to partisan voters only, that voters must preregister with their preferred party to vote in the primary, and even whether primaries are partisan at all (see McGhee et al. 2014 for a thorough and contemporary review of these regulations). Although many observers suspect that regulations on primary openness greatly influence the ideological extremity of the primary electorate, recent survey research finds that these regulations do little to affect the policy preferences of primary voters on average (Hill 2015).

and tests of the SPD are key empirical contributions in the following chapters. The sections that follow introduce key terms for understanding my critique of the existing research and my contribution to it in this project.

1.1.1 Key concept: policy ideology

If we had an ideal test of the SPD's implications, the policy preferences of partisan primary voters are an essential ingredient. Primary voters are one of the key constituencies that a candidate must please in the SPD view of primary elections. When partisan voters in a district are more conservative, the SPD claims that the candidate experiences a pressure to stake out a more conservative campaign position, especially in the primary. This section briefly discusses this project's terminology around voter ideology, the groups in the electorate for whom these concepts are at play, and how relate to other political science research.

When this project discusses voter "preferences" or voter "ideology," it specifically refers to a notion of *policy ideology*. An individual's policy ideology is a summary of their policy views in a left-right ideological space. Policy views are naturally complex and multidimensional, and it is possible for individuals to hold beliefs across policy areas that would strike many political scientists as being "ideologically inconsistent" (e.g. Campbell et al. 1960). Policy ideology distills this complexity into average tendencies; voters who hold a greater number of progressive preferences about policy are more ideologically progressive, and vice versa for voters with more conservative policy preferences. Voters who hold a mixture of progressive and conservative beliefs are ideologically moderate.

Policy ideology is different from policy *mood*, since mood measures voter preferences for the government to do more or less than an ever-shifting baseline, while ideology meant to be directly comparable using only issue information (Enns and Koch 2013; McGann 2014; Stimson 1991). Policy ideology is thus a similar concept to any method that measures a hidden ideological summary from one-off issue-based stimuli. This includes ideal point scores

for members of Congress, Supreme Court justices, and even individual citizens (Clinton, Jackman, and Rivers 2004; Martin and Quinn 2002; Poole and Rosenthal 1997; Tausanovitch and Warshaw 2013; Treier and Hillygus 2009). Other researchers have called this concept “policy liberalism” (Caughey and Warshaw 2015), which orients the concept so that “larger” values represent “more liberalism.” For this project, I prefer to orient the construct as *policy conservatism*, which orients a scale so that larger/more conservative values correspond to “rightward” movements on a number line. I try to be conscious of the difference between *consistent* issue beliefs and *extreme* issue beliefs throughout this project. Consistently conservative issue beliefs do not necessarily imply that an actor is “extremely” conservative (Fiorina, Abrams, and Pope 2005b), and an actor may appear “moderate” even if they hold a mixture of non-moderate progressive and conservative issue beliefs (Broockman 2016).

This project views policy ideology in a measurement modeling context, which we return to in Chapter 2. Policy ideology affects voters’ issue beliefs, and while issue beliefs can be measured using a survey, policy ideology itself is not observable. Instead, policy ideology exists in a latent space, an survey items on specific issues reveal only limited information about voters’ locations in the latent space. This is different from summarizing policy views by adding or averaging policy responses, which implicitly assumes that all items about all issues are equally informative about ideology. Modern measurement approaches relax this assumption, instead viewing survey items as sources of correlated measurement error across respondents, leading to more careful modeling approaches for estimating a latent signal from noisy survey data (Ansolabehere, Rodden, and Snyder 2008). Following this modeling tradition, I refer to an individual’s location in policy-ideological space as their “ideal point,” the point at which their expected utility of a policy is maximized with respect to their ideological preferences.

1.1.2 Key concept: district-party groups

I argue that another key construct at work in the SPD is the notion of *groups* in the electorate. For a given district, the general election is a contest among all voters, so we consider this constituency as a group. We sometimes refer to this group as the “general election constituency,” since it contains anybody who is eligible to vote in the general election. It does not specifically refer to voters only, but contains any citizen who could potentially be a voter in the general election. This ambiguity of who among the general election constituency actually votes is important to understanding a candidate’s incentives during the campaign, since the candidate is uncertain whether certain campaign tactics will galvanize some constituents while alienating others.

Another important grouping for this the partisan constituency within a district. Each congressional district contains constituents who are aligned with the Democratic Party or the Republican Party. I call these two groups of constituents *district-party groups*. All 435 congressional districts contain voters from the two major parties, totaling 870 district-party groups. For brevity, I sometimes refer to district-party groups as “party groups” or “partisan groups.” A district-party group contains any voting-eligible citizen who resides in a given district and identifies with a given party. As with the general election constituency, membership in a party group is no guarantee that the constituent votes either in the primary or in the general election. The important fact is that they are nominally aligned with one party’s voter base over the other. As I discuss below, decomposing a district’s voters into separate party groups is the key theoretical innovation in this project. To the best of my knowledge, an empirical study of primary representation that decomposes the voter preferences into district-party groups has never been done, even though it is crucial for testing the implications of the SPD theory.

One important distinction about district-party groups is that they are made of constituents,

not organizations. For this reason, it is sometimes helpful to refer to district-party groups as district party “publics,” which emphasizes that the groups are composed of ordinary citizens (Caughey and Warshaw 2018). There is no formal registration requirement to be a member of a party group, only a partisan identification. This construction of district-party publics aligns most closely with Key’s “party in the electorate” rather than “party as organization” (Key 1955). This distinguishes party publics from interest groups, policy groups, “intense policy demanders,” or the “extended party network,” which are concepts that describe organizations or maneuvers by political elites rather than rank-and-file constituents (Cohen et al. 2009; Koger, Masket, and Noel 2009; Masket 2009). Although recent research has underscored the importance of elite actors in shaping party nominations, this project focuses specifically on testing the SPD, which is a voter-centric view of primary representation. We bring in important concepts from elite-driven stories of primaries as they apply to particular claims being tested in later chapters.

1.1.3 Key concept: district-party ideology

It is important to define both “policy ideology” and “district-party publics” because they combine to form a key concept that anchors the substantive contributions of this project. This concept is *district-party ideology*: policy ideology aggregated to the level of the district-party group. Just as any individual might have a policy ideology ideal point, and any individual might affiliate with a party, district-party ideology averages the ideological variation within a district-party group into one group-level ideal point. By aggregating policy ideology within groups in this way, this project summarizes how policy ideology differs between Democrats and Republicans in the same district, and it shows how Democratic and Republican party groups vary across congressional districts. This enables us to consider how candidates are responsive to partisan sub-constituencies that together make up a shared general election constituencies (see also Clinton 2006).

1.1.4 Key concept: candidate campaign positioning

As with individual voters, we can imagine that candidates for Congress have campaign platforms, or at least promises and stated issue positions, that are located in ideological space as well. The study of United States politics most commonly places elite political actors in ideological space using their voting records, including members of Congress, Supreme Court justices, federal judges, and state legislators (Clinton, Jackman, and Rivers 2004; Epstein et al. 2007; Martin and Quinn 2002; Poole and Rosenthal 1997; Shor and McCarty 2011a). Researchers have extended the modeling intuitions to estimate ideal points from unconventional sources of data, including surveys of congressional candidates, campaign finance transactions, interest group ratings, text from political advertisements, and even Twitter activity (Ansolabehere, Snyder, and Stewart 2001; Barberá 2015; Bonica 2013; Burden 2004; Burden, Caldeira, and Groseclose 2000; Henderson 2016).

This project is interested in the ideological locations of candidates for office as measured through their campaigns. The positioning of campaigns is more directly related to the strategic positioning dilemma than any other concepts that we might scale in ideological space: candidates compete against one another by positioning themselves to appeal to a partisan base of voters, and partisan constituents consult use these campaign positions to nominate the candidate of their liking. To be sure, campaign positions are influenced by other activities that researchers have used to scale candidates for office. Incumbent legislators cast votes to form a defensible record in office, for instance, which both bolsters and constrains their campaign messages (Canes-Wrone, Brady, and Cogan 2002; Mayhew 1974). Not every primary candidate has a roll-call voting record to compare, however, so this project requires an ideal point measure that places incumbents, candidates challenging incumbents, and candidates running for open seats in a comparable ideological space.

This project measures primary candidates' campaign positioning using CF scores from

Bonica's (2019b) *Database on Ideology, Money in Politics, and Elections* (DIME) database. CF scores use campaign contributions to measure the political ideologies of contributors and recipients of campaign contributions. Because a wide variety of political actors engage in the contribution and receipt of campaign funds, the DIME contains CF score estimates for political candidates, party organizations, PACs, and individual donors. Unlike interest group ratings, another source of ideology scores for non-incumbent political candidates, CF scores are not constructed with a political agenda that implicitly "weights" issues according to the interest group's priorities (Fowler 1982; Snyder Jr 1992). CF scores assume that a donor makes financial contributions to political actors to maximize their utility over all potential contribution they could make, where utility decreases with greater ideological distance between donors and candidates. CF scores are the estimated ideal points for contributors and recipients that maximize this utility (Bonica 2013, 2014). These scores have been used in other studies of primary candidate ideology by Thomsen (2014), Thomsen (2020), Rogowski and Langella (2015), Ahler, Citrin, and Lenz (2016), and Porter and Treul (2020), and similar donation-based ideal point measures by Hall and Snyder (2015) have been used by Hall (2015) and Hall and Thompson (2018). As I discuss in future chapters, CF score are not without controversy as indicators of elite ideology, especially when comparing members of the same party (Hill and Huber 2017; Tausanovitch and Warshaw 2017), but other research shows that donors differentiate moderate and ideological candidates within the same party (Barber, Canes-Wrone, and Thrower 2016), the ideology component of CF scores outperforms a party-only model of giving (Bonica 2014), and CF scores predict future votes by members of Congress to a similar degree of accuracy as roll-call based scores do (Bonica 2019a).

1.1.5 The strategic positioning dilemma, implications, and research questions

Now that we have defined some key terms, we can see how they relate to previous research on the strategic positioning dilemma. The theory states that candidates balance two competing

constituencies during their campaign for office. Candidates face incentives to cater to the median voter in the general election, but they do not progress to the general election without first catering to partisan voters in the primary election. As a result, their campaign position is tailored to split the difference between the two constituencies, perhaps leaning more to the partisan base in safe districts and to the median voter in competitive districts. This section unpacks this intuition in detail and argues that existing research does not test the key claims.

First, how does district-party ideology affect the way candidates position themselves in a campaign? The logic of the SPD suggests that, at minimum, district-party conservatism should be positively correlated to the conservatism of a candidate's campaign position. At maximum, more conservative partisan voters exert a positive causal effect on the conservatism of a candidate's campaign position. This implies that candidates can perceive the conservatism of their partisan constituents, reflecting the relative variation in actual constituents' views if not the absolute level (Broockman and Skovron 2018).

Second, if candidates anticipate partisan voters' policy views and position themselves accordingly, this suggests that candidates believe partisan voters are capable of voting in accordance with their policy views. If this is true, we should expect that district-party groups that are more conservative should be more likely to nominate conservative nominees in primary elections.

These two predictions are the core empirical implications of the “strategic positioning dilemma” theory of representation in primaries. Crucially, testing each prediction requires a researcher to observe the policy ideologies of partisan constituents within a district, which is a separate group from the general election constituency or the location of the median voter. This project argues that district-party policy preferences are either absent from existing research or thoroughly misconstrued—an important theoretical and methodological point that I unpack in Section 1.2.3. As a result, U.S. elections research has been unable to empirically evaluate a widely held theory of representation in primaries.

Stated differently, this dissertation asks if primaries “work” the way the SPD claims they do. It is widely believed that primaries are effective means for voters to inject their sincere preferences into the selection of candidates and, in turn, the priorities of elected officials. Is this *actually* true? The two empirical research questions underlying this project are:

1. Do candidates position themselves to win the favor of primary voters?
2. Do primary voters select the candidate who best represents their issue beliefs?

1.2 Does the Strategic Positioning Dilemma Describe Primary Representation?

1.2.1 Theoretical concerns

The strategic positioning dilemma view of U.S. primaries has reasonable intuitions, but there are reasons to doubt some of its theoretical premises. First, the SPD is put forth as a theory to explain divergent candidate platforms across parties, but there are numerous theories that explain candidate divergence that do not rely on bottom-up pressures from primary voters. And second, the SPD requires voters and candidates to be highly sophisticated actors. Candidates must be capable of perceiving the relative extremity of their constituents, and voters capable of learning about candidate platforms, differentiating between candidates, and acting on sincerely-held preferences over candidate platforms.

The notion of the SPD emerges from a clash between idealized candidate positioning in formal models and the candidate positioning we observe in the real world. Classic formal models highlight a strategic logic for candidates to position themselves by “converging” to the location of the median voter: if constituents vote primarily with policy-based or ideological considerations, then candidates maximize the probability of electoral victory by positioning

themselves as closely to the median constituent as possible (Black 1948; Downs 1957).²

Empirical work finds evidence in partial support of both convergent and divergent candidate incentives. Candidates who run in electorally competitive districts are more moderate than co-partisans who are running in districts that run in electorally “safe” districts (Ansolabehere, Snyder, and Stewart 2001; Burden 2004), and even candidates who run in safe districts are marginally rewarded for taking more moderate issue positions than a typical party member would (Canes-Wrone, Brady, and Cogan 2002). Extremist candidates, meanwhile, earn fewer votes and are less likely to win in Congressional elections, and this tendency is stronger in competitive districts than in safe districts (Hall 2015). Despite these incentives to take moderate campaign positions, candidates nonetheless take divergent rather than convergent stances by and large. Republican and Democratic members of Congress vote very differently from one another, and this partisan divergence increased in recent years (McCarty and Poole 2006; Poole and Rosenthal 1997). The difference in legislative voting behavior across parties isn’t simply because Republicans and Democrats represent different districts, since Republicans and Democrats who represent similar districts (or the same state, in the case of U.S. Senators) nonetheless vote differently from one another (Brunell 2006; Brunell, Grofman, and Merrill 2016; McCarty, Poole, and Rosenthal 2009). Even among Congressional races in the exact same district, there is a sizable gap between Republican and Democratic candidate positions (Rogowski and Langella 2015). And although qualitative evidence from decades past suggests that candidates take careful positions on issues of local concern (Fenno 1978), more recent systematic tests find mixed evidence of localized, particu-

²Some empirical studies of candidate positioning (e.g. Ansolabehere, Snyder, and Stewart 2001; Brady, Han, and Pope 2007) claim that these formal models “predict” candidate convergence at the median voter. In my opinion, this misrepresents the formal work. Downs (1957) in particular explains the logic of candidate convergence, but he also explores many circumstances that would prevent the convergent equilibrium from appearing in the real world. This is important to clarify because, although it is common to describe candidate convergence as a “Downsian result” or a “Downsian prediction,” we should recognize that the convergent equilibrium is an oversimplification. Understanding the theoretical incentives that promote candidate moderation is more important than the whether we observe perfect candidate convergence empirically.

laristic position-taking. (Canes-Wrone, Minozzi, and Reveley 2011; Fowler and Hall 2016). In total, even though there is some evidence that candidates benefit by positioning themselves as marginally more moderate or more in line with local public opinion, the dominant finding is that candidates take divergent positions that are more closely aligned with a national party platform than with a set of local issue priorities.

The Downsian logic is a strong “centripetal” force that promotes moderation among candidates, but what “centrifugal” forces explain the non-moderate stances (Cox 1990)? Political scientists have explored several theories whose underlying mechanisms are distinct from the SPD notion of competing constituencies. Parties are interested in cultivating long-term reputations for pursuing certain policy priorities (Downs 1957; Stokes 1963). It benefits both major parties for these reputations to be distinct from one another, since parties have office-seeking motivations to mutually divide districts into geographic bases that tend to support one party platform consistently over time (Snyder 1994). Party leaders maintain these party reputations by constructing brand-consistent legislative agendas and pressuring legislators to support reputation-boosting legislation (Butler and Powell 2014; Cox and McCubbins 2005; Lebo, McGlynn, and Koger 2007). In turn, non-median party platforms are more appealing to constituents with ideologically consistent issue beliefs. Candidates benefit by rewarding these constituents in particular because they are more likely to influence election outcomes in favor of the candidate (Hirano and Ting 2015). These voters are more likely to turn out in general elections than moderate voters are, so it is more efficient for candidates to cater to these constituents. Partisan constituents are also more likely to engage in pro-party activism, such as staffing campaigns, contributing financially to campaigns, and attending party conventions (Aldrich 1983; Barber 2016; La Raja and Schaffner 2015; Layman et al. 2010; McClosky, Hoffmann, and O’Hara 1960).

These incentives for candidates to diverge from median positions are possible without considering primary elections whatsoever. Even if we introduce primary elections into the

theoretical story, many plausible explanations for divergence do not rely on outward pressures from ideological primary voters either. Many scholars of political parties maintain that parties retained their gatekeeping roles over party nominations even as the direct primary ostensibly removed their formal powers over candidate selection. Although primary campaigns take place, these scholars argue that an informal network of party actors wields enormous influence behind the scenes, controlling which candidates obtain access to the party's resources, donor lists, and partisan campaign labor (Cohen et al. 2009; Maskett 2009). Through these mechanisms, candidates can live or die by the nomination process long before primary *voters* ever enter the picture.

One reason to doubt the SPD on theoretical grounds is that it has high demands of voter sophistication in primary elections. It is well understood that learning about the characteristics and issue positions of political candidates is costly for voters, particularly in non-presidential elections. Party labels on the ballot are valuable heuristics for voters to differentiate the issue positions of Republican and Democratic candidates likely hold (Hill 2015). Primary elections, however, occur most of the time between candidates in the same party,³ which denies voters' the informational shortcut of a candidate's party affiliation (Norranders 1989). Primary elections often occur during months when voters are paying less attention to politics, and the press cover primary campaigns less closely than general election campaigns. Primary voters have a reputation for being more attentive and sophisticated consumers of political information, but in these lower-information environments, they may cast their ballots for non-policy reasons by prioritizing "Washington outsiders" or identity-based candidate features such as gender or race (Porter and Treul 2020; Thomsen 2020). They may also vote for the familiar candidate instead of the ideologically proximate one, in which case asymmetric campaign expenditures or news coverage may advantage one candidate over the other.

³There are a few exceptions to this institutional configuration of intra-party nominations. Some states hold blanket primaries, top-two primaries, or "jungle" primaries, where candidates from all parties compete on one ballot to be included in a runoff general election.

For example, Bonica (2020) attributes lawyers' numerical prominence in Congress to their ability to raise early money from their wealthy social networks. Furthermore, despite the disproportionate news coverage received by primary candidates who challenge incumbents on ideological grounds, the absolute number of explicitly ideological primary challenges in a given election cycle is low (Boatright 2013), so primary voters are unlikely to experience a deluge of policy-focused campaign messages even if they are attentive and sophisticated to receive and process those messages. In short, the claim that voters' policy preferences affect their choices in primary campaigns sounds straightforward, but the information environment of primary campaigns makes it difficult for constituents to vote foremost with their policy ideologies.

The SPD also requires candidates to perceive the policy ideologies of their partisan constituencies accurately in order to position their candidacies in relation to the partisan base and the median voter. Broockman and Skovron (2018) lend contradictory evidence to this notion by measuring the degree to which politicians "misperceive" their constituency's policy views. The authors find that elected politicians believe that their constituents are much more conservative on many issues than they actually are, which could affect how accurately candidates position themselves in relation to constituent views.

1.2.2 Empirical ambiguity

Empirical support for the strategic positioning dilemma is as unclear as the theoretical underpinning. When researchers conduct empirical tests of the SPD or the narrower premises of primary representation and competition on which it rests, the results are ambiguous and often contradictory of the SPD story. This section reviews existing research in this area to review the outstanding questions and preview the substantive innovations in this project.

Much of the interest in primary elections and representation comes from a focus on candidate divergence and partisan polarization. Why do candidates who stand for general

election take divergent stances from one another, and do the competitive dynamics of primary elections increase this divergence? Prominent studies of candidate positioning in general elections initially found conflicting evidence about the influence of stiff primary competition on candidate extremity. Using survey data from congressional candidates during the 2000 campaign, Burden (2004) finds that general election candidates take more extreme policy positions in their campaigns if they also faced stronger primary competition. This makes sense especially if primary candidates care more about the candidate's ideological positioning than general election voters do, the latter of whom are also receptive to non-policy appeals. Ansolabehere, Snyder, and Stewart (2001) find the reverse pattern using 1996 survey data. The gap between major party candidates was actually smaller when one of the candidates faced stiffer primary competition. This counter-intuitive finding makes sense if the presence of a primary challenger is itself a consequence of candidate positioning. If an incumbent maintains a partisan reputation, this may fend off credible primary challengers who have less room to wage an ideological campaign against the incumbent. As a result, the *threat* of a primary challenge exerts a centrifugal force on candidate positioning, even if a primary challenger never actually appears (Hacker, Pierson, and others 2005). Hirano et al. (2010) study this threat-based hypothesis by measuring potential primary threat as the average presence of primary competitors in down-ballot races. In district with high levels of latent primary threat, we might expect the incumbent to take more extreme stances in Congress. Although the idea that incumbents vote as party faithfuls to preempt opportunistic challengers is intuitive and supported by other research (e.g. Mann 1978), this measure was not meaningfully related to the extremity of an incumbent's voting record in Congress (Hirano et al. 2010). In short, the evidence of the polarizing effects of primary challenges is mixed and unclear.

Researchers interested in the polarizing effects of primaries on candidates and legislators has also examined primary "rules." Political parties are private organizations, and nominees are intended to represent the parties' priorities and governing values, but participation in

primary elections is not always restricted to party members only. Primary “openness” rules that govern who can participate in a partisan primary are managed by state election law, with some allowances for parties to set rules within those limits. States with “closed” primaries restrict participation in primaries only to individuals who are registered as Republicans or Democrats in their state registration records. States that allow third-party or non-partisan voters to participate in partisan primaries are “partially” open, and states where any voter can participate in any primary are regarded as “open” primaries. I discuss finer details of primary rules in later chapters. Researchers seeking to exploit state-level variation in primary rules hypothesize that states with more restrictive participation criteria might select more ideologically extreme primary nominees, and states with more relaxed rules might select relatively moderate nominees. This is because primary voters are commonly believed to hold more ideologically consistent policy views than other constituents, so candidate polarization will respond to the polarization among the voting public (Jacobson 2012). However, the consensus among recent studies finds little evidence supporting the hypothesis that primary rules affect polarization in congress or candidate divergence more broadly. This is because there is little consensus in public opinion research that partisans who participate in primaries are much different from partisans who do not participate in primaries, either demographically or ideologically (Geer 1988; Hill 2015; Jacobson 2012; Norrander 1989; Sides et al. 2018), though these studies cover many years, and the dynamics of primary voting might have changed. And even recent studies that find that primary voters hold more ideologically consistent views find no evidence that closed primaries nominate candidates that are more ideologically off-center (Hill 2015). This finding appears to hold for the House, Senate, and state legislatures through the past several decades (Hirano et al. 2010; McGhee et al. 2014; Rogowski and Langella 2015). Even reforms that drastically change the primary rules, such as California’s recent shift to a blanket primary where candidates in all parties compete for the same limited number of positions on the general election ballot, do not nominate legislators

whose voting records are much more moderate than before (Bullock and Clinton 2011).

These studies are incomplete in important ways that bear on the key substantive questions underlying this project. Most of these studies evaluate primaries' effects on representation by examining roll-call votes only. Since roll-call votes are only observable for incumbents, many of these analyses cannot measure candidate *divergence* because they cannot compare incumbents to non-incumbents nor two open-seat candidates. Some notable studies examine non-incumbent candidates for general election using candidate surveys (Ansolabehere, Snyder, and Stewart 2001; Burden 2004), but these studies are also limited because they do not observe the positions of candidates who lose the primary nomination. Without observing primary losers, we have no way of knowing if the general election candidate was relatively moderate or ideological in comparison to other primary candidates. It is much rarer for a study to measure primary candidate positioning as the key outcome variable using a method that covers incumbents, challengers, and open-seat candidates (Rogowski and Langella 2015).

1.2.3 Vote shares do not identify policy ideology

Another important drawback of the existing research on primaries and ideological representation is the way these studies handle voters' policy preferences. The strategic positioning dilemma pits two constituencies in a district against each other: the nominating constituency (district-party group) that contains constituents from one party's base, and the general election constituency that contains constituents from both major parties and with no party affiliation. The former is theorized to prefer ideologically faithful candidates who adhere closely to a partisan policy platform, while the latter prefers moderate candidates in the general election. Studies routinely acknowledge this distinction in theory, but they often abandon the distinction between the two groups in applied studies, instead operationalizing the preferences of all three constituencies—the general constituency and two partisan primary constituencies—using the same measure: the district-level presidential vote.

This project argues that the presidential vote is not a suitable for the study of primary representation for the simple reason that votes are not equivalent to policy preferences or policy ideology. Votes are choices that voters make under constraints, namely, the distance between the voter and the presidential candidates. Even in simple models where ideology is the only factor influencing vote choice, observing a voter's choice of candidate contains very little information about their ideological location. In the aggregate, Republican voters in a district may be ideological moderates or ideological conservatives, and the fact that they vote Republican does not inform us on the ideological distribution of Republican voters. Similarly, a district's vote outcome captures how all of its constituents vote *on average*, but because partisans tend to vote foremost for their preferred party even in the face of strong policy disagreements with the candidate (e.g. Barber and Pope 2019), aggregate vote shares for a district could easily be more affected by the *number* of Republicans and Democrats in a district rather than the exact location of their ideological preferences. Using the terminology by Tomz and Van Houweling (2008), studying vote shares rarely presents a "critical test" of theories of voting because the same observable vote outcome can arise from many underlying voter preference configurations.

Stated differently, the observed vote share in a district does not uniquely identify any important features of the underlying preferences of voters. Figure 1.1 demonstrates the problem using a simple theoretical model of ideological voting for president. We begin by demonstrating the the basic mechanics of the scenario in the two left-side panels. In this scenario, we consider one congressional district that contains many constituents. Every constituent has a policy ideal point represented on the real number line, with larger values indicating greater policy conservatism. Every constituent also identifies with either the Republican Party or the Democratic Party. The top-left panel breaks voters into Democratic and Republican Party affiliations and shows the probability distribution of ideal points within each partisan base, which in this example are both Normal distributions with a scale of 1.

Republican-identifying constituents hold policy preferences that are more conservative than Democratic constituents on average: the median Republican and Democrat are respectively located at 1 and -1.⁴ There is enough within-party variation that some Democratic constituents are more conservative than some Republican constituents, despite their party affiliation. The bottom-left panel combines the two partisan distributions into one distribution for the entire constituency. We assume at first that both partisan constituencies are equally sized, so the composite distribution is a simple finite mixture of the two distributions.⁵ The midpoint between two presidential candidates is shown at policy location 0. Assuming all constituents vote according to single-peaked and symmetric utility functions over policy space, constituents are indifferent between candidates if they have ideal points equal to 0, vote for the Democratic candidate if they have ideal points less than 0 (shown in darker gray), and vote for the Republican candidate if they have ideal points greater than 0 (shown in lighter gray). The aggregate election result, therefore, is equal to the cumulative distribution function of the combined distribution evaluated at the candidate midpoint. In the bottom-left panel, the vote share for the Democrat is 50%, with some Democrats voting for the Republican candidate, and some Republicans voting for the Democratic candidate.

The panels on the right side of Figure 1.1 show how slight changes to one party's preference distribution affects the aggregate distribution of preferences in the combined constituency and, as a result, the presidential vote share in the district. The composite distribution is again shown in gray, with dark and light shades indicating vote choice as in the bottom-left panel. The underlying partisan distributions are outlined only with red and blue lines to reduce

⁴Because these are Normal distributions, the median and the mean are equivalent. I refer to the median instead of the mean because medians are more directly relevant to spatial models of voting.

⁵Analytically, if $f_p(x)$ is the probability density of ideal points x in party p , then the composite density $f_m(x)$ is a weighted sum of the component densities: $f_m(x) = \sum_p w_p f_p(x)$, where w_p is a mixture weight representing the proportion of the total distribution contributed by party p , with weights constrained to sum to 1. In this first example, both partisan constituencies are equally populous, so both parties have weight $w_p = \frac{1}{2}$. If parties had different population sizes within the same district, w_p would take values in proportion to those population sizes.

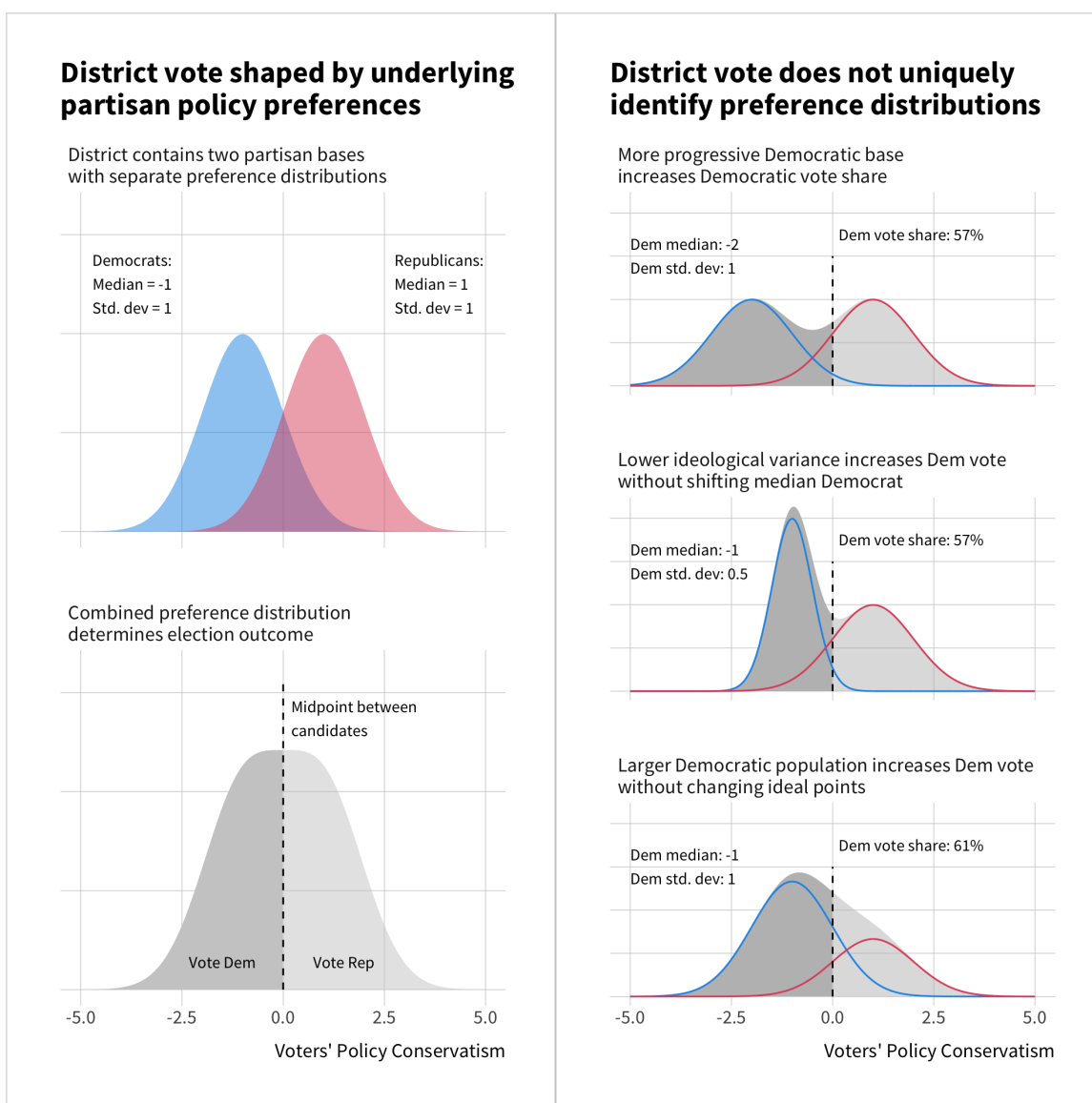


Figure 1.1: Demonstrating how district vote shares from a single election are insufficient to identify underlying policy-ideological features of the district. The left side shows how the policy preference distributions for two parties in a district (top panel) combine to form an aggregate preference distribution for the district as a whole (bottom panel). The right side shows how the Democratic vote share is affected by changes to either the locations, the scales, or the population sizes of the underlying partisan distributions.

visual clutter. The modifications to the underlying partisan preferences are simple, but even these simple changes reveal the fundamental problem with using district voting as a proxy for policy ideology in the voting population. In each panel, I intervene on only one feature of the Democratic Party ideal point distribution, leaving the Republican distribution untouched (median of 1, standard deviation of 1). Intervening on just one component of one party's distribution is meant to keep the demonstration simple, bearing in mind that the problem is much more complex in the real world, where we can imagine multiple simultaneous changes to both parties at once. The interventions highlight two classes of problems. First, we can perform multiple modifications of the underlying partisan distribution that obtain the same aggregate vote share. This proves that the district vote does not uniquely identify the characteristics of the underlying voter distributions. And second, we can alter the district vote outcome by changing party *sizes* without any change to the ideal point distribution within either party. This proves that vote shares may vary across districts even if partisan ideal points distributions are the same.

In the top-right panel, I shift the location of the Democratic ideal point distribution to the left, from a median of -1 to -2. This location shift results in a greater number of Democratic constituents with ideal points left of the candidate midpoint, increasing the Democratic vote share in the district from 50% to 57%. In the middle-right panel, I shrink the scale of the Democratic ideal point distribution from a standard deviation of 1 to a standard deviation of 0.5. Lower ideal point variance within the Democratic base has the exact same effect on the vote as shifting the location: more Democratic voters left of the midpoint, which increases the Democratic vote share to 57%. This means that compared to a district with a 50% presidential vote split, we would not be able to attribute the increased Democratic vote to a constituency that is *more progressive on average* (location) or simply *less heterogeneous* in its policy preferences (scale). The bottom-right panel in the figure shows how we obtain a different district vote without changing the underlying ideological distribution in either

party whatsoever, instead changing only the relative population size of each partisan base. The Democratic base in the final panel is unchanged compare to the original distribution laid out in the top-left: median of -1 and standard deviation of 1. The only difference is that the district contains an unequal balance of partisan voters, two Democratic constituents to every one Republican constituent. This results in an increased Democratic vote from 50% to 61%—ironically, the largest impact on the overall district vote despite not changing the ideological distribution of either party.

To review the lessons of Figure 1.1, observing a Democratic vote share greater than 50% reveals very little about the underlying distribution of voters. In every panel, we observe an increase in the Democratic vote compared to our baseline scenario, but the the median voter in either party does not need to change in order for vote shares to be affected. Since the Republican distribution is identical in every panel, inferring that Republicans are less conservative in districts with greater Democratic voting would be incorrect in every case. For the Democratic constituents, inferring a more progressive Democratic median voter from greater Democratic voting would be wrong in two of the three cases.

It is worth repeating that the scenario laid out in Figure 1.1 is a vast oversimplification of the real electorate. This is intentional, as it shows how intractable the problem becomes even in an artificial setting where we can take many variables as given. This scenario contains no complicating elements such as non-partisan or third-party identifiers, non-policy voting, random sources of utility or utility function heterogeneity across different voters, differential turnout between partisan bases, and so on, that we might incorporate directly into a formal model. It also does not take into account the inconveniences of real election data, where short-term forces impose additional shocks to vote shares that are unrelated to underlying voter preferences.

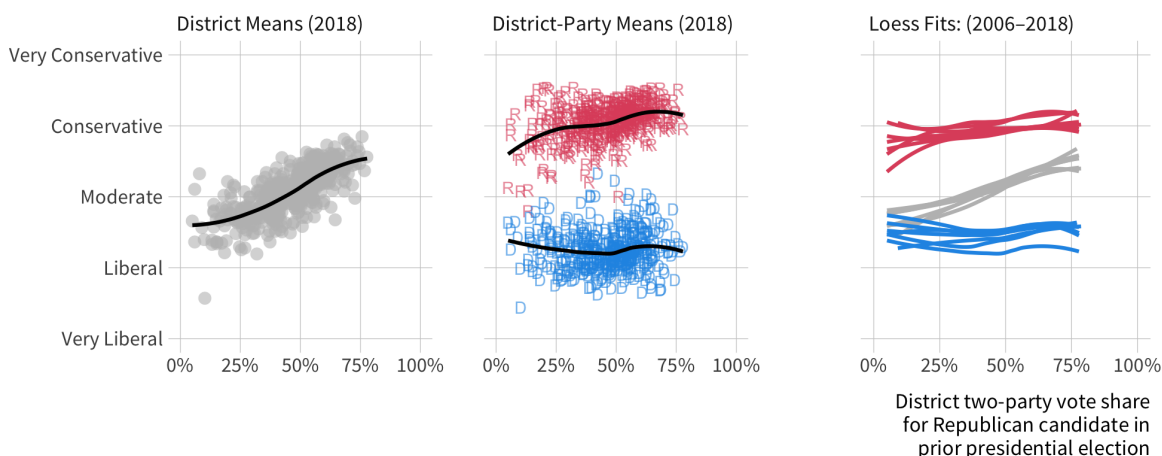
The conceptual difference between district vote shares and aggregate ideology appears in real data as well, as shown in Figure 1.2. The figure shows ideological self-placement responses

to the Cooperative Congressional Election Study (CCES) as an approximate measure of a citizen policy ideology. I calculate the average self-placement for all respondents in each congressional district, as well as the average self-placement of Republican and Democratic identifiers as separate subgroups within each district. The first two panels use 2018 data to show that the district vote captures variation in ideological self-placement reasonably well when examining congressional districts as a whole, but it does a poorer job capturing variation in self-placement within each party. The first panel shows that districts that voted more strongly for Democratic presidential candidate in 2016 were more liberal on average, and districts that voted more strongly for the Republican candidate in were more conservative, indicated by a positively sloped loess fit line. The middle panel shows that this pattern does not hold as strongly within parties. Among Republican identifiers within each district, a weaker but still positive relationship holds overall, with more conservative Republicans in districts that voted more Republican. Among Democratic identifiers, however, ideological self-placement is not as strongly related to aggregate voting, with a loess fit that is flatter and even negative at several points. The final panel of loess fits is included to show that this pattern appears in all CCES years and is not particular to 2018 CCES responses: a strong relationship between vote shares and self-placement *on average*, and weak or non-relationships within each party.

The substantive takeaway from Figure 1.2 is further evidence that we should doubt the use of aggregate voting in a district is a reliable proxy of ideological variation within partisan primaries. Because the presidential candidates are the same in each district in each year, we know that this mismatch isn't due to different candidates with different campaign positions in each district. Instead, the observed pattern suggests that any aggregate relationship between ideological self-placement and district voting is driven at least in part by the partisan *composition* of a district—more Republicans or more Democrats—rather than cross-district ideological variation within either party. As a result, studies that use the presidential vote

Weak Relationship Between District Voting and Ideology Within Parties

Average ideological self-placement in each congressional district



Data: Cooperative Congressional Election Studies

Figure 1.2: Average ideological self placement (vertical axis) and Republican vote share (horizontal axis) in all 435 congressional districts. Mean self-placement is calculated by numerically coding CCES ideological self-placement responses before averaging. The first panel plots average self-placement among all CCES respondents in each congressional districts. The middle panel breaks respondents in each congressional district into Republican and Democratic subgroups before averaging. The final panel plots loess fits for the same relationship measured over all CCES years.

to proxy within-party ideology may simply be measuring the *size* of a partisan group in a district instead of its ideological makeup.

Some researchers have recognized the identifiability problems with district presidential vote shares as a measure of district preferences. Levendusky, Pope, and Jackman (2008) specify a Bayesian structural model to subtract short-term forces on election results and isolate latent partisanship. Kernell (2009) formally proves that using a single election to cardinality place district ideal point medians is never possible, but that estimating the mean and variance of ideal point distributions is possible under distributional assumptions and a formal model of voting. Although these methods are promising innovations over the

common practice of using votes as a proxy for policy preferences, I have uncovered no studies of primary representation in the intervening years that have incorporated these methods. Furthermore, the methods estimate the median policy preference for a district as a whole. They do not describe separate partisan constituencies within a district, which is the essential missing ingredient.

I stress that this measurement problem is more than methodological nitpicking. The theoretical consequences are systemic. The literature's dependence on the presidential vote as a proxy for district preferences has prevented scholars from incorporating key theoretical constructs into empirical studies of primaries: the ideological preferences of partisan voters. Without serviceable measures of partisan policy preferences, we can say very little about the role of primary elections in the broader democratic order of U.S. politics. This affects our knowledges of topics beyond party nominations as well. To study how politicians weigh the opinions of various subconstituencies, which the study of U.S. politics is obviously interested in (Bartels 2009; Clinton 2006; Cohen et al. 2009; Fenno 1978; Gilens and Page 2014; Grossman and Hopkins 2016; Phillips 1995; Pitkin 1967), research must be able to measure the policy preferences of subconstituencies directly. The technology to estimate subconstituency preferences using survey data is admittedly quite new, and this district intends to continue this effort by extending existing models, highlighting important methodological considerations for model building and computation, and demonstrating how to use these measures for observational causal inference.

1.3 Project Outline and Contributions

1.3.1 Measuring district-party ideology

This chapter has so far identified a shortcoming in the study of primaries that subconstituency preferences are rarely measured. This project rectifies this shortcoming by measuring district-

party ideology for Republican and Democratic party groups in Chapter 2. This allows the project to carry out direct tests of SPD hypotheses that were previously impossible in Chapters 4 and 5.

I estimate district-party ideology this using an item response theory (IRT) approach to ideal point modeling. The model estimates the policy ideology for a typical Democrat and a typical Republican in each congressional district over time. I employ recent innovations in hierarchical modeling to measure individual traits at subnational units of aggregation using geographic and temporal smoothing (Caughey and Warshaw 2015; Lax and Phillips 2009; Pacheco 2011; Park, Gelman, and Bafumi 2004; Tausanovitch and Warshaw 2013; Warshaw and Rodden 2012). The model I build extends these technologies by specifying a more complete hierarchical structure for the bespoke parties-within-districts data context, a more flexible predictive model for geographic smoothing, and advances in Bayesian modeling best-practices from beyond the boundaries of political science (see also Section 1.3.5).

1.3.2 Empirical tests: how district-party ideology matters

After estimating the ideal point model for district-party groups, I apply these estimates in two critical tests of the strategic positioning dilemma.

Chapter 4 studies how district-party ideology affects candidate positioning in primary elections. If the primary constituency exerts a meaningful centrifugal force on candidate positioning, we should expect candidates with more ideological partisan constituencies to take more ideological stances, all else equal.

Chapter 5 studies how district-party ideology affects candidate selection in primary elections. If the primary constituency exerts a credible threat against candidates taking overly moderate campaign positions, we should expect more ideological constituencies to select more ideological candidates, all else equal.

An important institutional factor at play in each of these empirical settings is the mod-

erating effects of primary openness. Past studies have explored whether primaries that are closed to non-partisan and cross-partisan participation lead to the election of more extreme party nominees. District-party ideology is missing from these studies, but it matters for our theoretical expectations about the effects of primary rules. For instance, we should not expect a relatively partisan constituency to nominate an extremist candidate solely because the primary is closed to non-party members. Past studies have either ignored ideological variation across districts or used unsuitable proxy measures that do not measure district-party ideology. Including primary rules in Chapters 4 and 5 will provide a more faithful test of the primary rules hypothesis.

This project is not rooting for or against the veracity of the strategic positioning dilemma as a model of primary representation. The theory is intuitive and reasonable in its predictions for rational elite behavior, but its assumptions about voter competence and its empirical track record are less supportive of the theory. I wish for the empirical components of this project to be theory *testing* rather than advocacy for or against an idea in current political thought.

1.3.3 Causal inference with structural models

The strategic positioning dilemma is a story about the causal effects of district-party ideology on candidate positioning and candidate selection. Testing the theory requires a serious engagement with causal inference methods. Unfortunately, the observational data at work are difficult to manipulate in support of causal claims. District-party ideology is not randomly assigned, so we require methods for identifying unconfounded variation by design or adjusting for confounding with careful modeling.

One inherent limitation of the district-party ideology estimates is that they come from a measurement model. The measurement model smooths estimates with a hierarchical regression, where partial pooling improves the estimate for one unit “borrowing information” from other units. This shrinks estimates toward one another, imposing correlations between

estimates that share a common cause. To leverage exogenous variation for design-based causal inference, this variation would likely have to come predominantly through exogenous shocks to raw survey data, which is challenging to conceive of considering that many surveys must be pooled to achieve feasible estimates at the district-party level.

Given these data limitations, this project turns to causal identification through a conditional independence assumption (Rubin 2005), also known as “selection on observables.” Although selection on observables is a common approach to quantitative research, many analyses are not careful about their modeling choices, controlling for variables that do not improve causal identification or using modeling approaches that impose fragile or implausible functional assumptions on the data. One guiding ethic for the methodological contributions in this project is to take observational causal modeling more seriously than the existing research on primary representation by setting up empirical analyses that aspire to do the following:

- clearly state the potential outcomes model that links treatments, outcomes, and confounders.
- clearly state the causal estimand implied by a causal structure.
- clearly state the assumptions required to identify estimands and how modeling approaches relate to identification assumptions.
- use modeling approaches that are flexible enough to absorb confounding effects without too much dependence on strict functional forms.

I hope to satisfy these aims by invoking more explicit causal models of potential outcomes (Rubin 2005) and using “structural causal models” (SCMs) to guide model specification choices (e.g. Pearl 1995). The SCM approach makes heavy use of causal diagrams, or “directed acyclic graphs” (DAGs), to visualize a causal structure and identify causal claims. Causal diagrams as heuristic devices for causal inference are not new to political science in general

(Gerring 2001), but combining causal diagrams with the formal exactitude of the current causal inference tradition is less common in political science. Furthermore, SCMs and causal diagrams are less common in the literature on primaries and representation, which has not been as explicit about causal assumptions and empirical designs, with some notable exceptions (Fowler and Hall 2016; Hall 2015).

This project’s approach to causal inference has two stand-out contributions to the study of primary representation that would be impossible but for this approach. First, Chapter 4 contains a detailed discussion of the causal effect of district party ideology on candidate positioning *as mediated by* aggregate district partisanship. I lay out the causal structure in causal graphs, discuss identification assumptions required to estimate the causal quantity of interest, and implement a sequential-*g* modeling approach to estimate it (Acharya, Blackwell, and Sen 2016). Chapter 5 explores flexible modeling with machine learning (ML) as a way to reduce dependence on fragile model assumptions. The chapter discusses *regularization-induced confounding*, a statistical bias in a treatment effect estimate that arises when regularized estimators, such as those used in common ML methods, under-correct for strong confounding by injecting too much shrinkage into a statistical model. I show how to correct this bias using Neyman orthogonalization, a two-stage modeling approach that de-biases causal estimates by reparameterizing the structural causal model (Hahn et al. 2018). Regularization-induced confounding is a serious problem for high-dimensional causal inference, but it has been discussed almost nowhere in political science (Ratkovic 2019).

Selection on observables is a fragile assumption for causal identification, which leads many researchers to speak in “scientific euphemisms” about causality instead of invoking explicit causal language (Hernán 2018). I adopt the position that this “taboo against explicit causal inference” is harmful to the larger aims of a research program because it obscures the dependence of research findings on causal assumptions, whose transparency is essential for credible causal inference, and leads work to be misinterpreted by future audiences who tend to

interpret findings as causal regardless of author intent (Grosz, Rohrer, and Thoemmes 2020). No study will ever prove the existence of a causal effect. Researchers should be transparent about causal assumptions so that future readers and researchers have clearer ideas about how to improve previous work. As such, this work will invoke causal language, highlight identification assumptions, and discuss threats to identification assumptions openly.

1.3.4 Bayesian causal modeling

Another important methodological contribution in its Bayesian approach to causal inference. The key independent variable of interest, district-party ideology, is estimated using a Bayesian measurement model. It is not observed exactly, but it is estimated up to a probability distribution. Using those estimates in subsequent analysis requires some accounting for the uncertainty in those estimates. I do this by propagating the Bayesian framework from the measurement model forward into the causal models. Operationally, this is done by taking the posterior distribution from the measurement model and using it as a prior distribution in subsequent models, recovering a joint probability distribution that captures uncertainty in causal effects and its relationship to the uncertainty in the underlying data.

Although the Bayesian view of causal inference is not new (Rubin 1978), it appears almost nowhere in political science. Political scientists occasionally use Bayesian technology for analytical or computational convenience (e.g. Horiuchi, Imai, and Taniguchi 2007; Carlson 2020; Ornstein and Duck-Mayr 2020; Ratkovic and Tingley 2017), but rarely are the epistemic contours of Bayesian analysis explicitly credited for adding value to a causal analysis (Green et al. 2016; in economics, see Meager 2019).

Chapter 3 explores a Bayesian approach to causal inference in political science at length. It lays out a probabilistic model of potential outcomes adapted from Rubin (1978) and discusses how to interpret causal inference research designs through a Bayesian updating framework. I give pragmatic guidance for thinking about priors and specifying Bayesian causal models,

and I demonstrate the modeling approach by replicating and extending a few published analyses in political science, noting where the Bayesian approach leads to different conclusions and interpretations about the findings.

I apply Bayesian approaches to causal modeling in Chapters 4 and 5 by combining multistage models into one posterior distribution, which is natural for applied Bayesian modeling where causal effects can be summarized by marginalizing over “design-stage” parameters (Liao and Zigler 2020). Bayesian estimation is also valuable in Chapter 5 to quantify uncertainty in machine learning methods. This is done using a Bayesian neural network model, which automatically penalizes model complexity using prior distributions and quantifies treatment effect uncertainty in the posterior distribution (Beck, King, and Zeng 2004; MacKay 1992).

1.3.5 Bayesian best practices

Another important contribution of the modeling exercise is the detailed discussion of Bayesian modeling and computational implementation it contains. Classic Bayesian texts for political and social sciences are written for an outdated computational landscape where Metropolis-Hastings and Gibbs sampling algorithms were state-of-the-art estimation approaches (Gill 2014; Jackman 2009). Recent years have seen rapid progress in the development and understanding of Hamiltonian Monte Carlo algorithms, which are faster, more statistically reliable, and easier to diagnose (Betancourt 2017, 2019; Duane et al. 1987; Neal 2012), but they also require renewed attention to the way researchers specify and implement Bayesian models (Betancourt and Girolami 2015; Bürkner and others 2017; Carpenter et al. 2016). Furthermore, this new generation of applied Bayesian modeling has updated best practices for specifying priors, modeling workflow, and model evaluation that (to my knowledge) have no precedent in the current political science awareness (Betancourt 2018; Gabry et al. 2019; Gelman, Simpson, and Betancourt 2017; Lewandowski, Kurowicka, and Joe 2009; Vehtari,

Gelman, and Gabry 2017; Vehtari et al. 2020). One contribution of this project is to highlight the evolving landscape for Bayesian thinking and Bayesian workflow, which has not received its due attention as a new generation of political scientists explores Bayesian analysis.

— 2 —

Hierarchical IRT Model for District-Party Ideology

2.1 Testing the Model with Simulated Data

2.2 Data Sources

Describe data

2.3 Model Results

The model was estimated using a remote server at the University of Wisconsin–Madison.¹ I generated posterior samples using MCMC on 5 Markov chains. Each chain was run for 2,000 iterations, divided into 1,000 warmup iterations to tune Stan’s adaptive HMC algorithm and 1,000 post-warmup iterations saved for analysis.² Following the advice of Link and Eaton (2011), I stored every post-warmup sample with no thinning of chains, resulting in a total of 5,000 samples per parameter across all chains.³

Here we can reference Figure 2.1.

¹A Linux server (“Linstat”) maintained by Social Science Computing Cooperative.

²The algorithm was initialized with an `adapt_delta` parameter of 0.9 and a `max_treedepth` of 15.

³The chains mix well and exhibit little autocorrelation, which is owed to the fact that Hamiltonian Monte Carlo algorithms are much more efficient at proposing transitions and thus exploring a parameter space.

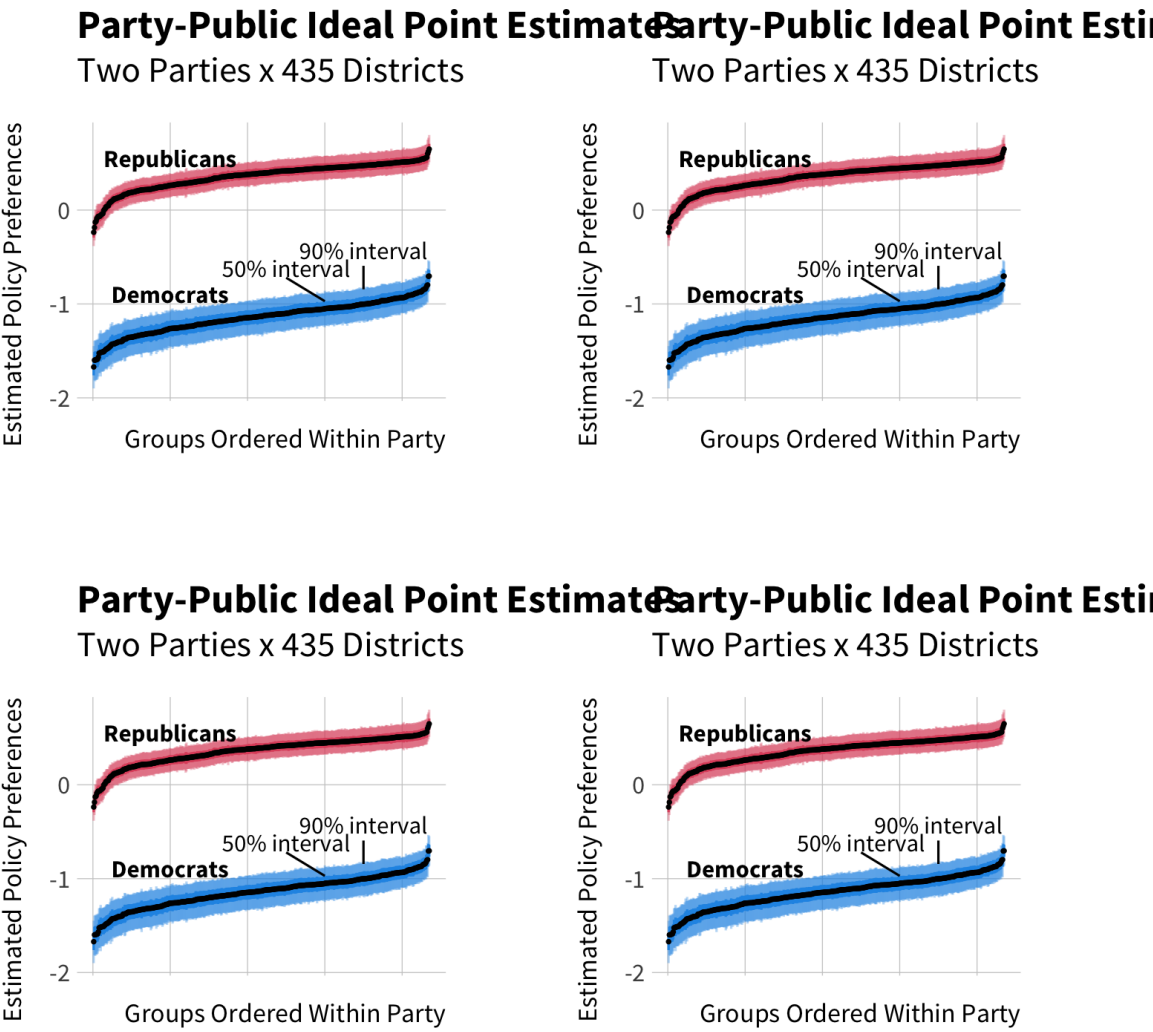


Figure 2.1: Posteriors

— 3 —

Bayesian Causal Inference

— 4 —

How District-Party Ideology Affects Primary Candidate

Positioning: A Bayesian De-Mediation Model

Do primary elections effectively transmit citizens' policy preferences into government? For this to be true, we should expect that the policy ideology with a partisan constituency to affect the ideological positioning of candidates who run for that party's nomination. This chapter explores the effect of district-party public ideology on the positioning of primary candidates running in that district.

It is important to distinguish the influence of the district-party public from the influence of the district overall. Does a candidate like Senator Susan Collins have a reputation as a moderate Republican because of a close balance between the number of Republican and Democratic voters in Maine? Or is the Republican constituency in Maine relatively moderate compared to Republican constituencies in states that elect more conservative Republicans? Although past research has been interested in the threat of primary challenges as a cause of ideological divergence between partisan legislators (for example Boatright 2013; Hill 2015; Hirano et al. 2010; McGhee et al. 2014), many of these studies lacked the capability to observe the preferences within local partisan groups as a concept distinct from aggregate partisanship

or aggregate voting in the entire district. This chapter uses my new measures of district-party ideology to investigate this question in ways that previous research projects could not.

The effect of district-party ideology on candidate positioning is a challenging causal inference problem. We cannot directly compare the “explanatory power” of district-party ideology and district-level voting by measuring whether one is more strongly correlated with candidate ideal points than the other, nor can we simply control for aggregate voting to recover the “partial effect” of district-party ideology. This is because aggregate policy ideology and aggregate voting are causally related: if a district contains a voter base with more conservative policy preferences, these policy preferences should influence aggregate voting behavior in the district as well as the positioning of candidates who try to respond to those policy preferences. Simply controlling for district voting in a regression will likely introduce collider bias by conditioning on a post-treatment variable (Greenland, Pearl, and Robins 1999; Montgomery, Nyhan, and Torres 2018).

This chapter advances this literature’s use of modern causal inference methods by estimating the effect of district-party ideology on primary candidate positions using sequential *g*-estimation, a structural modeling approach that measures the direct effect of district-party ideology while fixing the mediating effect of district-level voting. Substantively, I translate the primary candidate’s strategic positioning dilemma into the language of causal graphs, highlighting how aggregate districting voting mediates a relationship between district-party ideology and candidate positioning. Methodologically, I take the sequential *g* method as it appears in political science (Acharya, Blackwell, and Sen 2016) and embed it in a Bayesian framework. The Bayesian framework estimates all components of the structural model simultaneously, quantifying uncertainty in all model parameters in a single posterior distribution. This includes measurement uncertainty in ideal point estimates from the IRT model in Chapter 2, which is included as a prior distribution over ideal points. The key payoff for the Bayesian structure, therefore, is a unified framework for conducting inference about

treatment effects by marginalizing over other sources of uncertainty, including imprecise data and design parameters.

I find that primary candidates position themselves to fit district-party policy preferences: Republican candidates run more conservative campaigns in districts where the Republican constituency is more conservative, and Democrats run more progressive campaigns to please more progressive partisan constituencies. This finding holds even when controlling for aggregate district voting using sequential *g*-estimation. I also find, unlike other studies of primary representation, that primary candidates' responsiveness to district-party ideology is greater in closed primaries and weaker in open primaries.

4.1 Candidate Positioning and Voters' Policy Preferences

How do constituent preferences affect candidate positioning? This project explores the implications of what Brady, Han, and Pope (2007) call the “strategic positioning dilemma” (SPD), striking a balance between moderate position-taking to appease the general election constituency and ideological positioning taking to appease the partisan primary election constituency. Existing research contains plenty of studies that support the general theoretical intuition of the SPD theory, although I review some conflicts and ambiguities in detail in Chapter 1. To briefly review, general election candidates are rewarded at the ballot box for taking more moderate campaign stances (Canes-Wrone, Brady, and Cogan 2002; Hall 2015) and aligning themselves with local public opinion on specific issues (Canes-Wrone, Minozzi, and Reveley 2011; Fenno 1978; though see Fowler and Hall 2016). Nevertheless, no candidate makes it to the general election without first winning a primary nomination, where many scholars theorize that candidates benefit by taking more ideological positions that represent conventional views within the party. This could be a within-party Downsian incentive: the median primary voter is a ideological partisan with off-median policy preferences, so candi-

dates take more extreme positions to appeal to partisan constituency preferences (Aldrich 1983; Burden 2001). This is consistent with evidence from safe congressional districts, where candidates experience less general election threat and can more freely position their campaigns to target the primary electorate (Ansolabehere, Snyder, and Stewart 2001; Burden 2004). Pressures for primary candidates to take non-median stances may come through mechanisms unrelated to bottom-up voter pressures, instead reflecting candidates' need to organize committed staff and volunteers for their campaigns (Aldrich 2011; Layman et al. 2010; McClosky, Hoffmann, and O'Hara 1960), seek campaign funds from policy-seeking contributors (Barber 2016; Barber, Canes-Wrone, and Thrower 2016; La Raja and Schaffner 2015), or garner support from policy-demanding groups that control access to connections and resources to support candidates (Cohen et al. 2009; Koger, Masket, and Noel 2009; Masket 2009).

As I explained in more detail in Chapter 1, the explicit evidence about the SPD from primary elections themselves is surprisingly weak. This is mainly because most studies do not explicitly measure the ideological preferences of primary voters, instead using aggregate measures of voting that are not able to identify policy preferences (Kernell 2009). Furthermore, most aggregate-level measures of policy ideology do not differentiate between partisan constituencies in the same district (e.g. Tausanovitch and Warshaw 2013), which is important for understanding how candidates respond to their specific primary constituency. Without data that more closely resembles the theoretical story in question, studies that claim to demonstrate that candidates “handle the [strategic positioning] dilemma by positioning themselves closer to the primary electorate” rarely show anything of the sort (Brady, Han, and Pope 2007, 79). Brady, Han, and Pope (2007) show that among incumbent members of Congress, more liberal Democrats and more conservative Republicans attract fewer primary challenges and perform better in primary elections, but neither of these findings say anything specific lessons about how candidates are positioned relative to primary constituencies.

Hirano et al. (2010) get closer to this project's contribution by creating a statewide measure of primary electorate ideology from exit poll questions on ideological self-placement. They find that incumbent NOMINATE scores are more strongly related to general electorate ideological self-placement than primary electorate self-placement, a similar pattern as shown in Chapter 1, Figure 1.2. They further find no evidence that incumbent members of Congress do not have more extreme NOMINATE scores under greater threats of primary competition or higher primary turnout. This finding conflicts partially with Clinton (2006), who uses a large opinion survey of partisan voters in every congressional district to show that Republican members of Congress have NOMINATE scores that are more strongly related to the median Republican in their district, while Democrats' NOMINATE scores are more strongly related to the overall district median constituent. The IRT measures of district-party ideology that I create in Chapter 2 is a more direct measure of policy preferences than ideological self-placement scores or additive issue scores, which respectively tap a large amount of "symbolic" or identity-focused conceptions of political ideology or do not accord for differential measurement error across policy issues (Ellis and Stimson 2012; Treier and Hillygus 2009). My IRT ideal point scores, to my knowledge, are the first ideology scores for district-party groups to be applied to the study of congressional primaries.¹

Research on primary competition and candidate positioning is also held back by the availability of primary candidate data, which until recently was not available. Past studies of primary competition and polarization typically use ideology scores from legislative roll call votes, which include only incumbent members of congress or state legislators (Brady, Han, and Pope 2007; Hirano et al. 2010; McGhee et al. 2014), or they use surveys of general election candidates, which may include non-incumbent candidates for the general election but no candidates who ran in the primary and lost (Ansolabehere, Snyder, and Stewart 2001;

¹For IRT estimates of partisan constituencies and "ideological nationalization" at the U.S. state level, see Caughey, Dunham, and Warshaw (2018).

Burden 2004). Recent ideal point methods that use political financial contributions data have the capacity to scale a much broader universe of political actors, including candidates, political parties, PACS and interest groups, and donors (Bonica 2013, 2014; Hall and Snyder 2015). Few notable studies have yet used these contribution-based scores to study primary candidates (Ahler, Citrin, and Lenz 2016; Porter and Treul 2020; Rogowski and Langella 2015; Thomsen 2014, 2020).

Even with better measures of district-party ideology and candidate positions, there are some reasons to suspect that district-party ideology does not have a straightforward effect on candidate extremism. Thomsen (2014) finds that politicians are less likely to run for Congress (versus a state office) if their moderate stances make them a weaker “fit” for their party, measured as a greater distance between a candidate’s CF score and the average CF score in their state-party. This process could weaken the relationship between district-party ideology and candidate positioning because moderate candidates select themselves out of the campaign, even if they might appeal to district partisan voters. Porter and Treul (2020) find that the number of primary candidates who lack prior elected experience is increasing over time. Interestingly, they find no pattern between candidate experience and candidate extremism. This non-relationship may indicate that the positioning of inexperienced candidates may be less related to district-party ideology, either because the candidates are not adept at perceiving local ideology, they do not have the resources to conduct or acquire surveys of their constituencies, or emphasize non-ideological appeals in their campaigns. Incumbents, in turn, may be more responsive to local ideology as a result of their political skills and survivorship bias—ill-fitting candidates don’t become incumbents in the first place—although incumbents systematically misperceive public opinion as well (Broockman and Skovron 2018).

4.1.1 Primary rules

More recent studies of primary candidates tend not to focus on the overall relationship between voters' policy ideology and candidate positions. Instead, there is an enduring interest in the way primary rules affect candidates' positioning incentives by altering the composition of the primary electorate. Primary "rules" refer to regulations in state election law that control which voters can participate in primary elections and which inclusion criteria parties can define within those legal confines. A state primary is "closed" if only registered members of a political party are allowed to vote in the party's primary election. On the other side of the spectrum, an "open" primary allows any registered voter to cast a vote in any party's primary. There are many more states whose primary rules fall between these two extremes, such as "semi-closed" systems that allow party-unaffiliated voters to choose which party primary they want to vote for, even if registered partisans must vote in the primary of their party registration (McGhee et al. 2014). Political scientists and political observers speculate that these rules shape candidates' positioning incentives by changing the composition of primary voters. Closed primary elections, so the argument goes, are limited to registered party identifiers in a district, so candidates running in closed primaries must appeal to a more ideologically homogeneous primary electorate in order to win the nomination. More open systems, especially "blanket" systems where candidates from all parties run on the same primary ballot to advance to the general election, encourage primary candidates to take more moderate stances to attract a more ideologically diverse primary coalition. The general hypothesis, therefore, is that more restrictive primary rules exacerbate ideological polarization in congress and in state legislatures, and furthermore that polarization might be combated by moving primary elections to more open rules.

This "primary rules hypothesis" receives essentially zero support across several studies of U.S. elections. Hill (2015) studies the relationship between primary rules and the ideological

makeup of the primary electorate, asking if primary voters in states with more open primary rules are in fact more moderate on average than primary voters under closed primary rules. Estimating an IRT ideal point model on individual CCES respondents in each congressional district and validated voter turnout data, Hill finds that primary voters are more ideologically consistent than general election voters in the same party, but primary voters are no more extreme in closed primary states than in open primary states. Even if primary electorates are not affected by state primary rules, candidates may still suspect that primary rules matter and position themselves accordingly. Rogowski and Langella (2015) study the relationship between primary rules and candidate positioning as measured with CF scores. They find no systematic evidence that either congressional candidates or state legislative candidates are more extreme in closed primary states or more moderate in open primary states. McGhee et al. (2014) also study state legislators but using ideology scores that bridge the NOMINATE ideal point space to state legislative voting (Shor and McCarty 2011b), again finding no convincing evidence that primary systems matter for polarization in state legislatures. Within-state studies of changes to primary rules over time find mixed and highly qualified results, focused primarily on California's change to a "top two" primary system.² Bullock and Clinton (2011) find that the shift to the top-two primary promoted the election of more moderate candidates in California in competitive districts, measured using the two-party presidential vote, but no effects in more lopsided districts. Looking at the mechanisms underlying this, however, Hill (2015), who finds no effect on the ideological composition of primary voters after California's reforms, and survey experiments by Ahler, Citrin, and Lenz (2016) broadly show that voters are unable to identify which candidates are more ideological and which are more moderate.

I leverage my new data on district-party ideology to revisit the primary rules hypothesis in Section 4.3.1 below. Unlike most studies, I find that more candidates are less responsive to

²In a top-two system, candidates from all parties compete in a single primary for two spots on the general election ballot, which are awarded to the top two plurality winners in the primary.

district-party ideology in states with more open primary rules, consistent with the primary rules hypothesis.

4.1.2 Ideology within the two major parties

The SPD claims that candidates should be differentially responsive to primary and general electorates, but other theories on the ideological nature of U.S. parties may be relevant as well. An increasingly prominent theoretical perspective in U.S. political research holds that the parties are not asymmetrical but instead exhibit many “asymmetries” that help explain recent political conflict. In particular, the Republican Party is understood as an “ideological” party committed to a smaller welfare state, less regulation of business, and conservative cultural values, while the Democratic Party is a “group-based” party whose priorities reflect the mixture of social groups that compose the party’s core constituency (Grossman and Hopkins 2016). The mixture of group interests within the Democratic Party leads to internal conflicts about which policies to prioritize, while the Republican Party is more concerned with who is a “real Republican” or a “real conservative” (Freeman 1986). The ideological consensus in Republican political thought provides constituents with many different values-based rationales for supporting conservative policies, while Democrats remain more conflicted about how to rationalize their desires for activist government policies against an individualistic American value system that downplays the significance of group identities (Feldman and Zaller 1992; Free and Cantril 1967; Lelkes and Sniderman 2016).

The ideological foundations of U.S. political parties could be relevant for the way candidates conceive of their “responsiveness” to constituents. Because the Republican Party has an ideological underpinning, and elite political actors will be more aware of partisan ideology than individual constituents will be, Republican candidates may not exhibit much ideological responsiveness even if their constituents’ *policy attitudes* contain real variation. In other words, the ideological identity of the Republican Party could be a stronger organiz-

ing principle for Republican candidate positioning than the heterogeneous views of local constituencies. Democrats, meanwhile, may appear more responsive to district-party ideology because local opinion variation reflects the social group profile of the constituency, which is the organizing feature of Democratic Party representation. The intuition of these “asymmetric party” predictions diverge from (Clinton 2006), who finds that Republicans are *more* responsive to within-party opinion variation than Democrats but doesn’t provide much theoretical exploration of why this should be the case.

4.1.3 Exploratory analysis

The analysis begins by examining the topline correlation between district-party ideology and candidate positions. For the dependent variable, I use the dynamic CF score included in the DIME congressional candidate database for 2012, 2014, and 2016 candidates (Bonica 2019b). For district-party ideology, I use the mean from the MCMC samples of the IRT model in Chapter 2, which are estimated from polling data over the 2010s districting cycle. Using only the mean understates the amount of uncertainty in the ensuing analysis, which is later corrected in the full analysis. For now, these initial investigations serve only to give us an impression of the raw data.

Figure 4.1 shows the topline relationship between primary candidate CF scores and the ideal point mean for the district-party they ran to represent. Each point represents a primary candidate for Congress in either the Democratic or Republican Party primary in years 2012, 2014, and 2016 as they appear in the DIME congressional database. This totals 1,975 Democratic candidates and 2,197 Republican candidates over three election cycles. In addition to each candidate, I plot least-squares regression lines calculated separately for each party. Confidence intervals reflect standard errors that are clustered at the district-party level to capture correlated error among candidates who run in the same primary race.

The figure shows a weak but decisively positive relationship between ideal point means

Candidate Positioning and Group Ideology

Candidates from 2012, 2014, and 2016

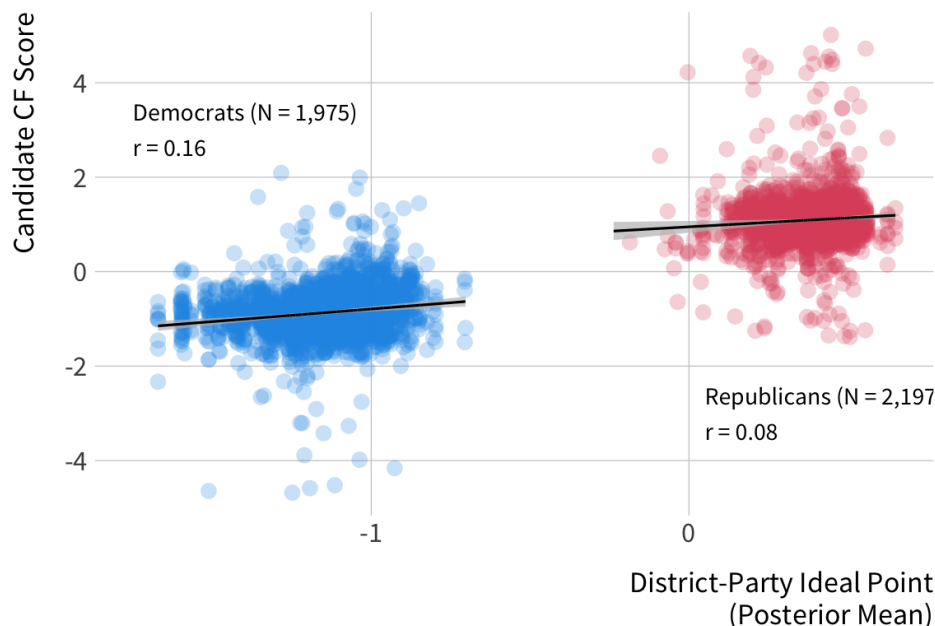


Figure 4.1: Topline relationship between district-party ideology and candidate positions. The horizontal axis plots the posterior mean for a group ideology, and the vertical axis is the dynamic CF score for primary candidates included in the DIME congressional candidate database.

and CF scores; district-parties that are more conservative see more conservative primary candidates. The relationship is stronger for Democrats than for Republicans in slope (0.53 versus 0.38) and in correlation (0.16 versus 0.08). Because one-unit increases are large on the ideal-point scale, it is helpful to standardize these coefficients in terms of standard deviations in the raw data. Conveniently for a bivariate regression, the standardized coefficient is equivalent to the correlation coefficient. Increasing district-party ideology by one within-party standard deviation is associated with a CF score increase of 0.16 standard deviations among Democrats ($p < .01$) and 0.08 standard deviations among Republicans ($p = 0.01$). A small number of outlier CF scores exist for each party, but these outliers are few compared to the approximately 2,000 observations in each party. The regression lines track the center of

each party's ideal point distribution, so these outliers do not appear to be large influences on the topline relationship.

Figure 4.2 plots the same relationship with candidates divided into incumbents, challengers of incumbents, and candidates running for a district with no incumbent running for reelection. It is immediately noticeable that CF scores have lower variance among incumbent candidates than among non-incumbent incumbents, and the correlations between CF scores and ideal point means are markedly higher: 0.39 among Democrats and 0.37. The overall higher correlation suggests incumbents could be more capable at positioning themselves for their primary electorates or that prior elections are effective at screening out ill-fitting candidates to represent the district. The higher correlation among Democrats also disappears among incumbent candidates, suggesting that no party is obviously more responsive to primary constituencies than the other, contrary to Clinton (2006)'s finding that Republican incumbents were uniquely sensitive to ideological variation within their partisan bases. CF scores are higher variance among challengers and open-seat candidates, and their relationship to district-party ideology is weaker but still positive. This may be because challengers and open seat candidates must seek other ways to appeal to candidates aside from ideological fit. This would be consistent with Boatright (2013)'s key finding that primary challenges focused purely on ideological fit to the district are relatively rare. A more mechanical explanation could be that CF scores are higher variance for challengers and open seat candidates because they struggle to raise from the same concentrated network of donors as incumbents do, attenuating the relationship between CF scores and district-party ideologies. Even still, the data generally suggest that candidates of all statuses have some awareness, on average, of how to position themselves as more conservative or progressive to corner their local partisan constituencies. Almost all relationships are statistical significant at a 1% level except among open seat candidates, whose sample sizes are also smaller.

Another notable finding among incumbents is the appearance of a much smaller "inter-

cept shift” between the two parties. While other studies typically find a large gap between Republicans and Democrats who represent otherwise equivalent districts [McCarty, Poole, and Rosenthal (2009); among others], the predicted CF scores for Democratic and Republican incumbents appear to diverge less dramatically if each regression line were extrapolated to meet at moderate values of district-party ideology. This interpretation comes with several caveats, naturally. First, there is no way to know if a linear extrapolation is an appropriate method for comparing parties in “otherwise equivalent districts,” since there are no districts whose partisan constituencies are similar enough to make that extrapolation without strict functional form assumptions. Second, a cursory regression analysis of CF scores on district-party ideology and party still finds mean difference between the two parties, even though it is smaller among incumbents. Nevertheless, the data broadly reinforce the theoretical notion responsiveness to partisan constituencies partially explains at least some of the ideological distance between Republican and Democratic candidates running in the same district. Future research on inter-party divergence could incorporate district-party ideology scores and address this issue more directly.

Incumbency Status and Ideological Responsiveness

Candidates from 2012, 2014, and 2016

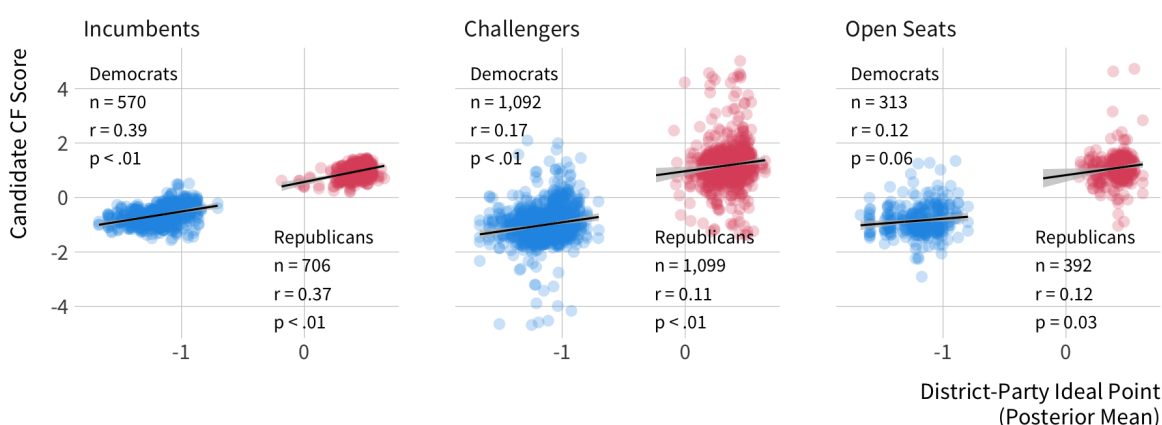


Figure 4.2: District-party ideology and candidate positioning across candidate incumbency status.

Figure 4.3 shows how the relationship between district-party ideology and CF scores varies by election year, splitting the 2012, 2014, and 2016 campaign cycles into three panels. There is no definitive trend toward increasing polarization or increasing responsiveness to partisan ideology in more recent years, which is inconsistent with theoretical speculation that candidates have become more sensitive to primary electorates over the past few elections. Instead, the same weak but positive relationship appears in nearly all subsets of data with no overwhelming explanation for differences over time or between parties. The flattest relationship actually appears most recently for Republicans in 2016, among whom the relationship is nearly flat. We should hesitate to interpret too much from one estimate, but future researchers could investigate whether financial contributions by Republican donors had a different ideological character in 2016, if perhaps Donald Trump's unusual campaign platform altered which donors wanted to support which candidates, or if moderate donors directed their money away from Republican candidates in anticipation of a Democratic national victory.

Ideological Responsiveness Across 2010s Districting Cycle

Each year contains incumbents, challengers, and open seats

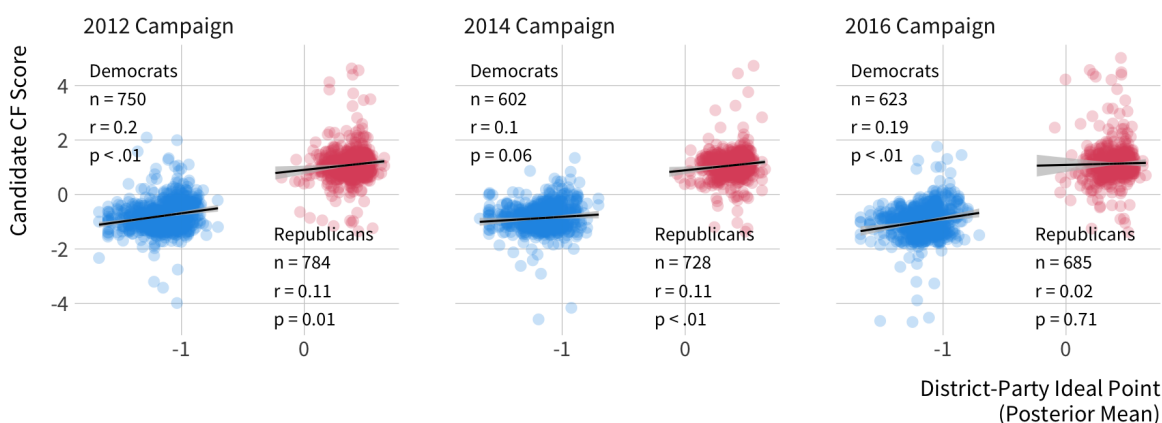


Figure 4.3: District-party ideology and candidate positioning across election cycles within the 2010s districting cycle.

Finally, Figures 4.4 and 4.5 plot the topline relationship between district-party ideology

and CF scores for candidates running in states with different levels of primary openness. Data on primary systems come from Boatright, Moscardelli, and Vickrey (2017) for 2012 and 2014, and I coded 2016 by consulting the National Conference of State Legislators, Ballotpedia, and OpenPrimaries.org.³ I code primary rules using a three-category scheme. “Closed” primaries allow only pre-registered partisans to participate in a primary election. “Semi-closed” primaries are closed to party registrants, but they allow independent or non-partisan voters to choose which primary in which they wish to participate. Lastly, “open” primaries allow any eligible voter to participate in any party’s primary. I code nonpartisan blanket and top-two primaries as “open.” I choose a coarser three-level scheme over the five-level scheme in McGhee et al. (2014) because it is unlikely that voters process the fine legal differences that lead the authors to classify (for instance) “semi-closed” states and “semi-open” states differently. The three-part scheme is also more specific than the two-part open/closed scheme used by Hill (2015), since it is difficult to group independent and non-partisan participation in semi-closed states as either closed or open.

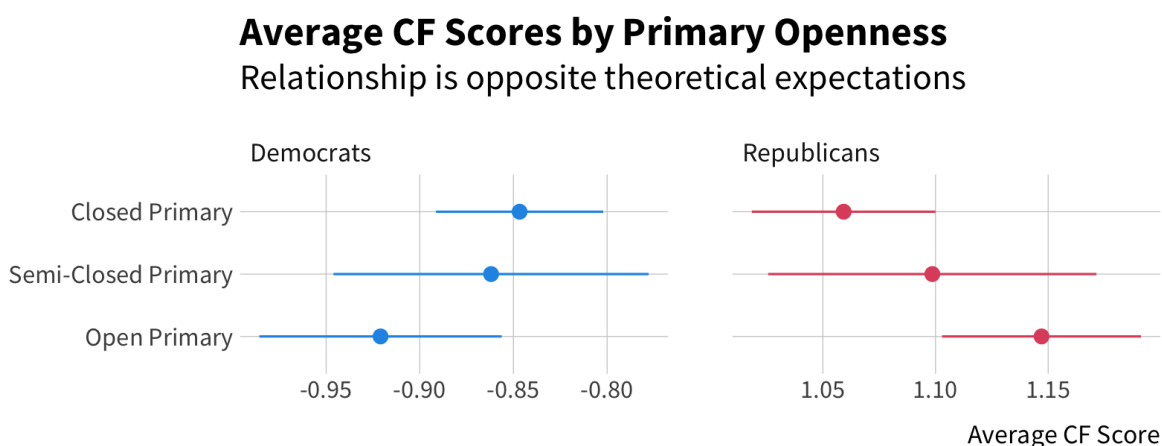


Figure 4.4: Average CF scores candidate positioning in states with closed, semi-closed, and open primary rules.

Conventional wisdom about primary openness implies that closed primaries should

³ Accessed May 27, 2020.

produce more extreme candidates, and open primaries should produce more moderate candidates. Figure 4.5, which plots the average CF score for Republicans and Democrats under each primary system, contradicts this hypothesis.⁴ Democrats are in fact more progressive and Republicans more conservative in states with increasingly open primary participation rules, which is the opposite direction of the commonly hypothesized pattern. Figure 4.5 goes on to plot the linear relationship between CF scores and district-party ideology in each state to see if candidates in closed primaries are more responsive to district-party ideology. The relationships do not support the primary rules hypothesis either. While Republican candidates are indeed least sensitive to local partisan ideology in states with the model open rules—indeed the point estimate of the relationship is negative—they are *most* sensitive in semi-closed rather than closed states. The evidence from Democrats is even less consistent with the primary rules hypothesis. Democrats in semi-closed systems are also most responsive to district-party ideology, and they are less sensitive to district-party ideology in closed primary systems than they are in open systems.

The descriptive results from Figures 4.1 through 4.5 inform a few modeling choices for the causal analysis to follow. Because incumbency status appears to modify the relationship between citizen and candidate ideology, some of the analysis below estimates the effect of district-party ideology using subsets of data on incumbents, challengers, and open seat candidates. By comparison, estimates exhibit no clear time variation, so I choose to pool election cycles into one model, using fixed effects where appropriate to the design, rather than estimating entirely separate models for different cycles. And although the descriptive relationships were broadly similar for both parties—contradicting an “asymmetric parties” prediction that Republicans would be less responsive to district party ideology as well as the Clinton (2006) finding that Republicans are *more* responsive— the variables that could

⁴Estimates are calculated from a linear regression on indicator variables for each primary system type with group-clustered standard errors.

Ideological Responsiveness Across Primary Rules

Collapsed coding of primary rules

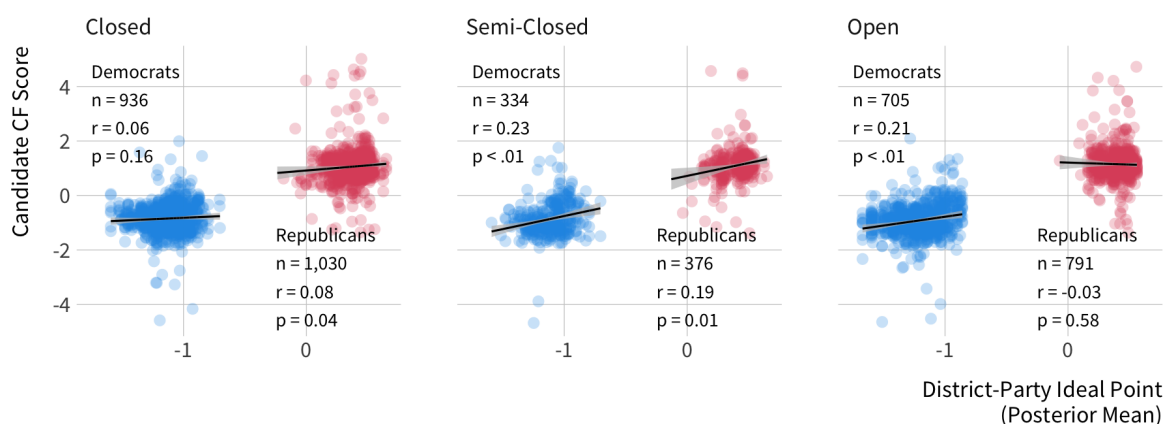


Figure 4.5: District-party ideology and candidate positioning in states with closed, semi-closed, and open primary rules.

confound the relationship between district-party ideology and candidate ideology are likely to differ dramatically across parties. In order to increase the credibility of the selection-on-observables regression design, I therefore estimate models for Democrats and Republicans separately.

4.2 The Causal Effect of District-Party Ideology

The descriptive picture in Figures 4.1 through 4.5 are suggestive about the relationship between district-party ideology and candidate ideology, but the correlational analysis is insufficient to identify the causal effect of the partisan constituency. As with many regression-based analyses using observational data, we are concerned about variables that confound the relationship between district-party ideology and candidate positioning—district characteristics that affect the degree of conservatism among both voters and candidates simultaneously. Even more troublesome than confounding is the causal structure linking theoretical components of the strategic positioning dilemma: how the ideological locations of the partisan constituency *and*

the total district constituency affect one another, and how they jointly influence a candidate's positioning calculus.

The strategic positioning dilemma (SPD) describes the campaign incentives that arise when a candidate chooses their campaign's ideological location for a multi-stage campaign season. Let CF_i represent the campaign position for candidate i in left-right ideological space. The theory states that candidates optimize their chances of winning by taking a position between the ideological location of their partisan constituency, denoted $\bar{\theta}_g$ for the district-party group g in which the candidate runs, and the ideological median among the district constituency, denoted V_d for the district d where group g resides.⁵ Figure 4.6 plots a hypothetical candidate's location between the partisan and district constituencies. Whether the candidate positions themselves closer to the partisan constituency or to the district constituency is a function of the candidate's perceived degree of electoral threat from each constituency. In an electorally safe district, the candidate of the advantaged party is reasonably assured to win the general election, so they may take a campaign position closer to the partisan constituency in order to neutralize the threat from other partisan candidates in the primary election. In a competitive district, the candidate faces greater general election competition, so they position themselves closer to the district constituency to avoid alienating moderate voters who could decide the election (Aldrich 1983; Burden 2001). As it relates to causal inference, the theoretical setup implies a causal effect of the primary constituency on candidate positioning (a causal path $\bar{\theta}_g \rightarrow CF_i$) and a causal effect of the district constituency on candidate positioning ($V_d \rightarrow CF_i$).

This analysis proposes an even more specific causal structure, the details of which are crucial for the research design and statistical approach. I invoke a causal model where the party constituency location affects both the district constituency location and the candidate

⁵We distinguish individual candidates i from district-party groups g and districts d , because every district contains two major party groups, and every group can contain multiple primary candidates in a given election cycle.

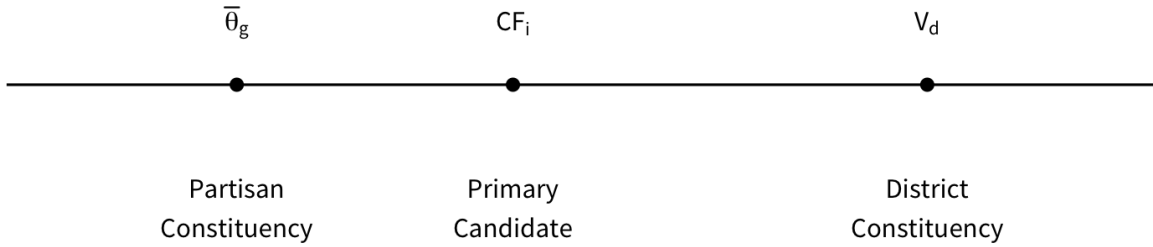
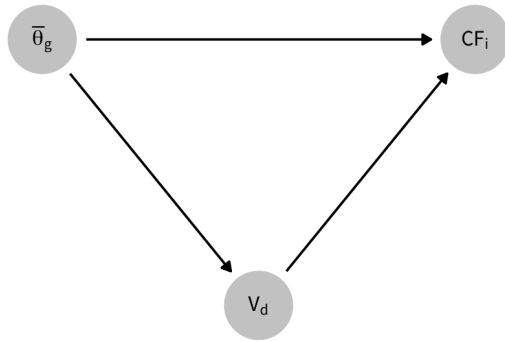


Figure 4.6: Spatial representation of key variables affecting the strategic positioning dilemma.

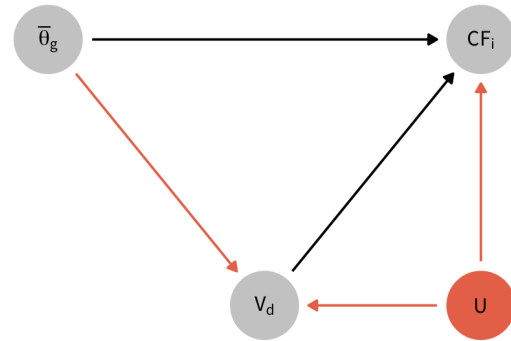
location, but the district location only affects the candidate location. The diagram in the left half of Figure 4.7 captures this structure. A causal structure where the party location affects the district location, but not the other way around, makes sense by appealing to the spatial theory underpinnings of the SPD. The party location and the district location are aggregations of many individual constituents who each have ideological ideal points. The district constituency contains all of the party constituency as well as any other constituent who has a different party affiliation or no party affiliation. This means that if we could causally intervene on the party constituency location by, for example, shifting it to the left, this entails a leftward shift among the individual partisans in that constituency. Because these partisan individuals are also members of the district constituency, this *ceteris paribus* intervention on the party location also directly affects the district location, justifying the causal path $\bar{\theta}_g \rightarrow V_d$. From a spatial theory point of view, the only way to intervene on a party location with no downstream effect on the district location would be to introduce an offsetting shift in the ideal points of constituents outside that party or an offsetting change in the partisan composition of constituents in district overall, neither of which is entailed by a causal intervention on the party location. Furthermore, intervening on the location of the *district* has no necessary effect on the party location. This is because the district location is affected by factors other than the partisan constituency, namely the ideal points of constituents outside that partisan constituency or the relative sizes of partisan constituencies within the district.⁶

⁶If we generalize the SPD to a panel data framework, we can stipulate a mechanism by which past district

SPD Causal Structure



Why Conditioning Fails



Controlling for V_d opens path through U

Figure 4.7: Left: A causal diagram for the strategic positioning dilemma. The district location V_d is a collider between district-party ideology θ_g and the candidate position CF_i . Right: Conditioning on post-treatment variables can bias causal estimates by creating artificial associations through unobserved confounders.

If our primary interest is the treatment effect of district-party ideology on candidate positioning, but the district location also affects candidate positioning, it is a natural impulse to want to condition on some measure of district preferences. For instance, the previous two-party presidential vote share is an available signal of district-level preferences that candidates can consult when they position themselves, so researchers may control for presidential voting in a district in order to isolate the effect of district-party ideology on candidate positioning. This is similar to the motivation of Clinton (2006), who compared incumbent responsiveness to district median preferences vs party median preferences by including measures of each variable in a regression. Because the district location is a post-treatment variable, or a “collider” (Pearl 2009) on the causal path from the party location to candidate location ($\theta_g \rightarrow V_d \rightarrow CF_i$), identifying the effect of district-party ideology separately from the overall district effect is

voting may affect future district-party ideology through a thermostatic opinion mechanism: heavy Democratic voting in one election leads to a Democratic presidency, which causes issue opinions across the country to become more conservative. The data in this project are unable to capture this mechanism because district-party ideology is measured for a district-party group over a districting cycle, but this is something that researchers could explore if they approach primary representation through a dynamic causal inference framework (Blackwell and Glynn 2018; Imai and Kim 2019).

more complicated than controlling for the presidential vote in a regression. This is because conditioning on a collider variable can bias causal effects by opening unblocked causal paths from treatment to outcome through unobserved confounders. This problem is diagrammed in the left half of Figure 4.7, which shows how conditioning on the presidential vote, a measure of V_d , opens a new pathway $\bar{\theta}_g \rightarrow V_d \rightarrow U \rightarrow CF_i$, biasing the estimated causal effect of $\bar{\theta}_g$. At the same time, doing nothing to adjust for aggregate district preferences may identify the total effect of the district-party constituency under certain assumptions, but it would not detect whether presidential voting absorbs all of the effect of the partisan constituency. There would be no way to contrast the unique impact of the partisan constituency from its downstream effect on overall district voting.

This analysis resolves this problem by using sequential g -estimation to identify a quantity called the average controlled direct effect (ACDE) of district-party ideology on candidate positioning (Acharya, Blackwell, and Sen 2016; Vansteelandt 2009). In terms of potential outcomes, let $CF_i(\bar{\theta}_g, V_d(\bar{\theta}_g))$ be candidate i 's CF score as a function of district-party ideology and the past presidential vote, which represents district-level preferences V_d and is itself a function of district-party ideology. The controlled direct effect (CDE) for a unit i is the difference in CF scores for different district-party ideology treatments, holding the presidential vote fixed at some value v .

$$CDE_i(\theta, \theta', m) = CF_i(\bar{\theta}_g = \theta, V_d = v) - CF_i(\bar{\theta}_g = \theta', V_d = v) \quad (4.1)$$

The ACDE is the average of the CDEs if all units were fixed at the same mediator value v .

Sequential g -estimation is a structural modeling routine that estimates ACDEs by subtracting intermediary causal effects without creating collider bias (Acharya, Blackwell, and Sen 2016; Vansteelandt 2009). The routine requires adjusting for two sets of confounders: pre-treatment confounders X_d that affect both district-party ideology and CF scores, and intermediate confounders Z_d that affect both the presidential vote and CF scores. Inter-

mediate confounders are allowed to be affected by both pre-treatment confounders and district-party ideology. The first graph in Figure 4.8 diagrams the stipulated causal structure among the district-party ideology treatment, the presidential vote mediator, and both sets of confounders. In the first stage of the model, the researcher estimates the effect of the mediator on the outcome variable, conditional on all confounders.

$$\begin{aligned} & \mathbb{E} \left[\text{CF}_i(\theta, v) - \text{CF}_i(\theta, v') \mid \bar{\theta}_g = \theta, X_d = x, Z_d = z \right] \\ &= \mathbb{E} \left[\text{CF}_i \mid V_d = v, \bar{\theta}_g = \theta, X_d = x, Z_d = z \right] \\ & - \mathbb{E} \left[\text{CF}_i \mid V_d = v', \bar{\theta}_g = \theta, X_d = x, Z_d = z \right] \end{aligned} \quad (4.2)$$

The left side of Equation (4.2) represents the conditional effect of the mediator in terms of potential outcomes, and the right side is the quantity that can be estimated from observed data assuming that the mediator is conditionally ignorable given all confounders and has positive assignment probability [Acharya, Blackwell, and Sen (2016);]. This specification blocks all back-door paths from V_d to CF_i , which can be seen in the first panel of Figure 4.8.

The next step of sequential g -estimation is to subtract the effect of the mediator from the outcome, also known as “demediation” or “blip-down.” Demediation removes all variation in the outcome variable that is attributable to the causal effect of mediator. This stage is algebraically equivalent to subtracting a *demediation function* from the observed outcome. The mediation function in terms of potential outcomes is as follows:

$$\delta_d(\theta, v, \bar{v}, x) = \mathbb{E} \left[\text{CF}_i(\theta, v) - \text{CF}_i(\theta, \bar{v} = 0.5) \mid X_d = x \right] \quad (4.3)$$

which represents the expected effect on CF scores by setting the presidential vote to its observed value versus some fixed reference value \bar{v} for all units, conditional on X_d (Acharya, Blackwell, and Sen 2016).⁷ In this analysis, I fix the two-party presidential vote to 0.5, an

⁷The model specification below functionally assumes to interactions between the presidential vote and intermediate confounders Z_d , so the specification of the demediation function in Equation (4.4) does not depend on Z_d , although this assumption is not strictly necessary for nonparametric identification of the ACDE (Robins 1997).

even split between Republicans and Democrats in the district. We subtract the demediation function from the original outcome to obtain the demediated CF score, $b(\text{CF})_i$, which is equivalent to the potential CF score if all units had a presidential vote of 0.5.

$$\begin{aligned} b(\text{CF})_i &= \text{CF}_i - \delta_d(\bar{\theta}_g, V_d, \bar{v} = 0.5, x) \\ \mathbb{E}[b(\text{CF})_i] &= \mathbb{E}[b(\text{CF})_i(\theta, V_d = \bar{v})] \end{aligned} \quad (4.4)$$

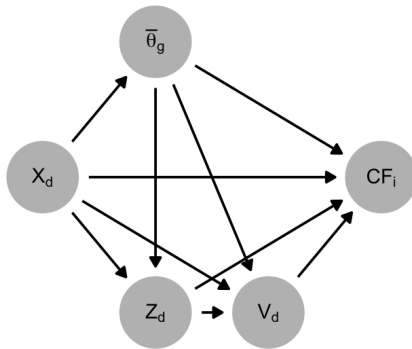
Finally, the researcher estimates the effect of the treatment on the demediated outcome, which is equivalent to the controlled direct effect on the original outcome.

$$\begin{aligned} &\mathbb{E}[\text{CF}_i(\theta, \bar{v}) - \text{CF}_i(\theta', \bar{v}) \mid X_d = x] \\ &= \mathbb{E}[b(\text{CF})_i(\theta) - b(\text{CF})_i(\theta') \mid X_d = x] \\ &= \mathbb{E}[b(\text{CF})_i \mid \bar{\theta}_g = \theta, x] - \mathbb{E}[b(\text{CF})_i \mid \bar{\theta}_g = \theta', x] \end{aligned} \quad (4.5)$$

The first statement in Equation (4.5) relates the controlled direct effect in terms of CF scores to the total effect on demediated CF scores. The second statement defines how the ACDE can be estimated with observable data under the assumption that assignment to district-party ideology ignorable and with positive probability given pre-treatment covariates (Acharya, Blackwell, and Sen 2016). The top-right graph in Figure 4.8 shows how the demediation step recovers the controlled direct effect of district-party ideology. Demediating the outcome removes all variation in CF scores caused by the presidential vote, so it deletes the path $V_d \rightarrow \text{CF}_i$ in the diagram. In turn, there is no need to condition on the presidential vote V_d in Equation (4.5) to identify the ACDE, even though the diagram shows that district-party ideology has an effect on the presidential vote. Furthermore, the graph contains remains a causal path from the intermediate confounders Z_d to the CF score, the stage-two model does not condition on these confounders because these pathways are a part of the district-party ideology's ACDE on CF scores. It is worth noting here that if we estimate the stage-two model using the original CF score rather than the demediated CF score, this would estimate the

Stage 1

Identifies mediator effect

**Stage 2**

Identifies controlled direct effect of treatment using demediated outcome

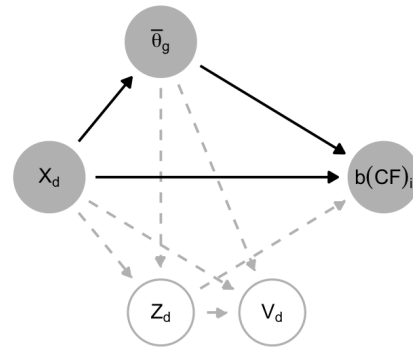
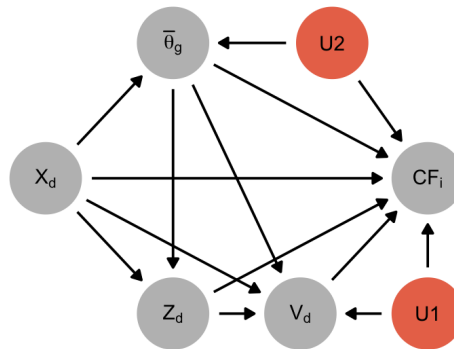
**Violations of Sequential Ignorability**In stage 1 ($U1$) and stage 2 ($U2$)

Figure 4.8: Causal graphs describing the modeling problem and sequential g estimation. The stage 1 graph identifies the effect of the past district voting (V_d) on candidate positioning (CF_i). The stage 2 treatment-outcome model subtracts the district vote effect from candidate positions and identifies the effect of district-party ideology $\bar{\theta}_g$ on the demediated CF score $b(CF)_i$, which is equivalent to the controlled direct effect on the raw CF score. The final graph shows where unadjusted confounders violate the nonparametric causal identification assumptions in stage 1 ($U1$) and in stage 2 ($U2$).

total effect of district-party ideology. This quantity can be valuable because the difference between the total effect and the controlled direct effect shows how much of the total effect flows is carried by a mediating mechanism.

The bottom graph in Figure 4.8 shows where unmeasured confounding can violate the two ignorability assumptions required to estimate the ACDE. The stage-one model identifies the causal effect of the past presidential vote, so if an unmeasured variable (represented in the figure by $U1$) affects both district voting and CF scores, the mediator's effect is not identified. Similarly, the stage-two model does not identify the effect of district party ideology on the demediated CF score if they share an unmeasured cause $U2$. Unmeasured variables in other locations of the graph certainly exist, but they do not violate the sequential ignorability assumptions unless they can be represented by an unblocked back-door path through $U1$ or $U2$.

With few exceptions, it is not typical for research on primary politics to incorporate explicit causal inference methodologies (Doherty, Dowling, and Miller 2019; Fowler and Hall 2016; Hall 2015; Hall and Thompson 2018). Although many other studies use a selection-on-observables research design to study candidate positioning in primary elections, a major contribution of this study's research design is the use of structural causal modeling to define, identify, and estimate a specific causal quantity.

4.2.1 Sequential g implementation

The average controlled direct effect of district-party ideology on primary candidate CF scores is *nonparametrically* identifying under SUTVA, sequential ignorability, and positivity. I estimate the sequential g model using a linear model specification laid out by Acharya, Blackwell, and Sen (2016). As with any linear model, this specification imposes additional functional form and distributional assumptions on the data, which are helpful for confounder adjustment in the absence of other design-based variation from instruments or discontinuities,

but could lead to incorrect inferences if assumptions are false. One goal for future work in this literature would be to combine the sequential g structural model with semi- and non-parametric estimation methods, an nascent area of work that exists largely outside of political science (Athey, Imbens, and Wager 2016; Chernozhukov et al. 2017; Hahn et al. 2020; Hill 2011; Ratkovic 2019; Samii, Paler, and Daly 2016; Wager and Athey 2018)

This section lays out the linear specification for sequential g -estimation. I define the model using notation that is general enough to apply to any subset of data where the model is estimated. As mentioned above, I investigate subsets of data by incumbency and primary rules, and I estimate each model separately for Republicans and Democrats. The data contain measures of candidates i in district-party groups g in districts d . Because all models are estimated separately for each party, groups and districts perfectly overlap. Nonetheless, I use g for variables that can vary across party within a district and d for variables that are fixed for both parties in a district. Because the estimation contains in two stages of regression modeling, I sometimes subscript some parameters with 1 or 2 to indicate which equation they belong to.

The first stage is a mediator–outcome model that estimates how the Republican two-party presidential vote in the past election V_d affects the CF score of candidate i within group g . We set up the sequential g method to control for the previous Republican presidential vote share in district d , denoted $pvote_g$. This is done with the following multilevel regression model.

$$\begin{aligned}
 CF_i &= a_0 + \mu V_{d[i]} + \eta \bar{\theta}_{g[i]} + \mathbf{x}_{d[i]}^\top \beta + \mathbf{z}_{d[i]}^\top \gamma + \alpha_{d[i]} + \varepsilon_i \\
 \alpha_d &\sim \text{Normal}(0, \sigma_\alpha) \\
 \varepsilon_i &\sim \text{Normal}(0, \sigma_\varepsilon)
 \end{aligned} \tag{4.6}$$

The group ideal point $\bar{\theta}_g$ is included in the regression as a control, so its coefficient η is estimated as a nuisance parameter, as are the coefficients β for district-level pre-treatment confounders \mathbf{x}_d , the coefficients γ for district-level intermediate confounders \mathbf{z}_d , and the

constant a_0 . Because the mediator is measured at the district level, I include a district error term α_d in addition to the candidate-level error term ε_i . This multilevel model accounts correlated error among candidates in the same district-party group, similar to clustering standard errors when a treatment is dosed at the cluster level.

We then use the estimates from the stage-1 model to demediate the CF score variable. Because the first stage is a linear model, the demediation function has a straightforward parametric definition,

$$\delta_d = \mu (V_d - \bar{v}) \quad (4.7)$$

where \bar{v} is the reference value for the mediator where all units are fixed, which I set equal to 0.5 to represent a 50-50 split in the previous two-party presidential vote.⁸ The demediation function is subscripted d because it varies across districts according to the observed value of the previous Republican vote share, which also entails that the demediation function is equivalent for all candidates in the same district-party group, regardless of their original CF score. We calculate the demediated outcome $b(\text{CF})_i$ by subtracting each district's demediation function from its observed outcome value,

$$b(\text{CF})_i = \text{CF}_i - \delta_d \quad (4.8)$$

The demediated outcome is then used in the second stage estimation of the controlled direct effect. The second stage is different from the first stage in two ways. First, because the demediated outcome fixes the value of the mediator, there is no more variation across observations that can be attributed to the mediator, so V_d is omitted from the equation. Second, intermediate confounders are omitted because they can be affected by district-party ideology, and as such should be left unadjusted in order to identify the ACDE. This new

⁸The demediation function can be more complex if the mediator's effect on the outcome is modeled with a more complex model containing interactions or nonlinearities.

equation is,

$$\begin{aligned}
 b(\text{CF})_i &= a_1 + \tau \bar{\theta}_g + \mathbf{x}_{g[i]}^\tau \omega + v_{d[i]} + u_i \\
 v_d &\sim \text{Normal}(0, \sigma_v) \\
 u_i &\sim \text{Normal}(0, \sigma_u)
 \end{aligned} \tag{4.9}$$

where a_1 is a constant, τ is the coefficient for district-party ideology, ω are coefficients for district-level pre-treatment confounders \mathbf{x}_d , v_d is a district error term, and u_i is a candidate error term. In a linear model specification, τ measures the average total effect of a one-unit increase in district-party ideology on demediated CF scores, which is equivalent to the ACDE on the original CF scores. More generally, the ACDE of setting district-party ideology from $\bar{\theta}'$ to $\bar{\theta}$ is as follows.⁹

$$ACDE(\tau, \bar{\theta}, \bar{\theta}') = \tau(\bar{\theta} - \bar{\theta}') \tag{4.10}$$

4.2.2 Bayesian modeling and interpretation

Another key methodological innovation in this chapter is embedding sequential g -estimation in a Bayesian framework. Chapter 3 describes a number of special advantages for Bayesian causal modeling, stemming from the fact that Bayesian inference allows the researcher to conduct posterior inference about causal effects using probabilities: the treatment effect is *probably* greater than x , where “probably” is defined in relation to an empirical cumulative probability distribution function of posterior samples.

The most important feature of Bayesian causal inference for this chapter is the fact that one posterior distribution quantifies uncertainty in all parameters from the multi-stage sequential g method. This is valuable because uncertainty in stages 1 and 2 are directly related to one another by way of the demediation function. Whereas non-Bayesian analysis requires either

⁹Again, the exact formula for the ACDE depends on the model specification. The linear model with no treatment interactions produces a simple linear ACDE definition, but a more complex model would entail a more complex formula.

ad-hoc variance corrections for the multistage model or a bootstrapping approach (Acharya, Blackwell, and Sen 2016), the posterior distribution from the Bayesian model captures all variances and covariances among model parameters by its very nature. Inferences about the ACDE can then be expressed by marginalizing the posterior distribution with respect to these auxiliary model parameters, isolating the remaining dimension of the posterior corresponding to the ACDE. Letting π represent all auxiliary model parameters, and using the definition of the ACDE in Equation (4.10), the posterior distribution for the ACDE is given by

$$p(\tau(\bar{\theta} - \bar{\theta}') \mid \mathbf{CF}) = \int p(\tau(\bar{\theta} - \bar{\theta}'), \pi \mid \mathbf{CF}) d\pi, \quad (4.11)$$

which is a the distribution of ACDE values that condition on the observed data and marginalize over the associated auxiliary parameter values. In practice, we use MCMC samples to approximate this posterior distribution by picking any two values of $\bar{\theta}$ and $\bar{\theta}'$, extracting posterior samples for τ , and calculating the ACDE for each MCMC sample iteration. Posterior expectations for the ACDE can be calculated, up to Monte Carlo error, by averaging the ACDE values from a draw of MCMC samples. Uncertainty intervals for the ACDE can be estimated by noting which MCMC samples bound the inner 90 or 95% of posterior samples.

The joint posterior distribution is also essential for incorporating uncertainty in district-party ideal points themselves. The IRT model in Chapter 2 does not estimate district-party ideal points exactly. Instead, ideal points are estimated only up to a posterior distribution, with uncertainty that reflects both prior ignorance and a finite sample of polling data. To estimate the effect of district-party ideology on candidate positioning, the causal analysis incorporates measurement error in ideal points “the Bayesian way,” where posterior uncertainty from one analysis becomes prior uncertainty in later analyses. One way to implement this “full uncertainty” would be to build one joint model containing both the IRT measurement model and causal inferential models, although this would be a burdensome feat of computer

programming and computationally expensive to run each model. Instead, I approximate the joint model by constructing a prior for $\bar{\theta}_g$ in the causal model by approximating the posterior distribution from the IRT model. This prior is constructed by defining the group ideal point $\bar{\theta}_g$ as an element of Θ , the vector of all group ideal points, which gets a multivariate Normal prior.

$$\Theta \sim \text{MultiNormal}(\hat{\Theta}, \hat{\Sigma}_{\Theta}) \quad (4.12)$$

The hyperparameters in the prior are estimated from MCMC samples for the ideal points. The mean vector $\hat{\Theta}$ contains MCMC means for each ideal point, and the matrix $\hat{\Sigma}_{\Theta}$ contains variances for each ideal point on the diagonal and covariances between any two ideal points on the off-diagonals. Because the IRT model partially pools ideal points using a hierarchical Normal regression, the multivariate Normal prior is a reasonable stand-in for representing prior ideal point uncertainty in causal analyses.

Including ideal point uncertainty as a prior distribution effectively adds a measurement error model on top to the sequential g analysis. Although Acharya, Blackwell, and Sen (2016) describe a variance estimator and a bootstrap method for dealing with multi-stage modeling uncertainty, neither of these methods is naturally suited to a measurement error context where an additional layer of ideal point uncertainty is represented in MCMC samples. Past research has explored the use of inverse-variance weighting to downweight observations with greater ideal point uncertainty, which accomplishes a similar goal as the prior distribution but throws away all prior information about the covariance between ideal points. More recently, researchers have employed numerical methods for “uncertainty propagation,” where the researcher estimates an uncertain quantity, simulates values from the quantity’s posterior distribution, and pushes those simulated values through a downstream analysis. Quantities of interest are then averaged across the posterior simulations (see Kastle et al. 2015, 791; Caughey and Warshaw 2018, 6, 2019, 360). This is similar in spirit to “multiple overimputation”

(Blackwell, Honaker, and King 2017), which extends multiple imputation to a measurement error setting by iteratively replacing mismeasured observations with draws from their posterior distribution. This project regards these propagation methods as insufficiently Bayesian because they “cut” the flow of information between models. Posterior cutting allows model 1 to inform model 2, but model 2 can never inform model 1 (Plummer 2015). This can be an undesirable modeling property if the causal model can inform the measurement model (Treier and Jackman 2008). For instance, if ideal points are related to candidate positioning, and ideal points are measured with error, then observing candidate positions can update our information about ideal points. By specifying the prior directly and updating the parameters, there is no need for any additional imputation steps or post-estimation model averaging (Gelman and Hill 2006, 542).

The Bayesian sequential g approach, of course, requires priors for other model parameters as well. As I describe below in Section 4.2.4, most variables in the model are binary indicators or are standardized to be mean 0 and variance 1. Furthermore, most of the outcome data in each model fall mostly in the $[-2, 2]$, so we can place standard Normal priors on the covariates without being overly informative about the covariate effects and introducing regularization bias (Hahn et al. 2018). Constants are given wider priors to account for the fact that not all predictors are exactly centered at means.

$$\begin{aligned}
 a_0, a_1 &\sim \text{Normal}(0, 5) \\
 \eta, \mu, \beta, \gamma &\sim \text{Normal}(0, 1) \\
 \tau, \omega &\sim \text{Normal}(0, 1)
 \end{aligned} \tag{4.13}$$

Both modeling stages contain Normal district-level error terms with estimated variances that facilitate partial pooling. These variances are themselves given half-Cauchy priors with weakly informative scale parameter values that are about half the range of the raw outcome

data within each party.

$$\begin{aligned}
 \alpha_d &\sim \text{Normal}(0, \sigma_\alpha) \\
 \sigma_\alpha &\sim \text{Half-Cauchy}(0, 1) \\
 \nu_d &\sim \text{Normal}(0, \sigma_\nu) \\
 \sigma_\nu &\sim \text{Half-Cauchy}(0, 1)
 \end{aligned} \tag{4.14}$$

And finally, each stage of the model has a Normal error term for candidates within districts. The variances for these errors are given half-Cauchy priors with wider scale values than the districts errors, since residual variation between any two candidates is likely larger than the variation between average candidates in any two districts.

$$\begin{aligned}
 \varepsilon_i &\sim \text{Normal}(0, \sigma_\varepsilon) \\
 \sigma_\varepsilon &\sim \text{Half-Cauchy}(0, 2) \\
 u_i &\sim \text{Normal}(0, \sigma_u) \\
 \sigma_u &\sim \text{Half-Cauchy}(0, 2)
 \end{aligned} \tag{4.15}$$

4.2.3 Causal inference with multilevel data

The research question in this chapter presents us with multilevel data: how is the ideological positioning of primary candidates affected by the policy ideology of partisans in their district, when there are potentially multiple candidates per district? In this scenario, the outcome is a variable specific to an individual candidate i , but the treatment is fixed for an entire district-partisan group g . This introduces a few issues for statistical assumptions and causal assumptions.

On the statistical front, multilevel models bias coefficient estimates when the aggregate errors are not exchangeable. Mechanically, this is similar to “omitted variable bias” in a single-level regression. Although this concern is well founded for many multilevel models, for these

models we can be less concerned. Because all predictors in these regressions are measured at the district level, the district error term is analogous to an error term that we would obtain by averaging every candidate's CF score within a district-party group and running a single-level regression on those averages. Both of these model specifications require an exchangeable errors assumption at the district level. The only difference for the models in this analysis is the additional candidate-level errors, but this too is a non-issue. Averaging candidate data within each district would invoke a similar assumption about the exchangeability of candidates given the district, otherwise it would be inappropriate to average data within a district.

Even though the multilevel model has similar assumptions as a regression on averages, it has certain benefits that are convenient for these data. Because the number of candidates in a district isn't fixed across all districts, we would expect heteroskedasticity in a regression-on-averages model, since some districts would have higher variances due to fewer candidates. In the extreme case, if a district contained only one unopposed primary candidate, a naïve estimator would be unable to distinguish district-level variance from candidate-level variance. The multilevel model addresses this by estimating the distributions of district errors and candidate errors simultaneously, enhancing the model's ability to recognize when larger district errors are caused by signal versus noise. Errors from smaller districts borrow more information from the overall distribution of districts, downweighting the contributions of smaller districts by shrinking their error terms toward a mean of zero. This has a similar intuition as a weighted least squares regression on the district-averaged data, where groups with more observations are more informative about global parameters and receive greater weight. This is yet another example where priors stabilize pathological model behavior, underscoring the flexibility afforded by Bayesian model-building for confronting the idiosyncrasies of a dataset with tactics that are both intuitive and feasible.

The multilevel data structure also raises causal inference issues that are worth clarifying. As with many causal models where treatments are assigned to clusters of observations, it

makes sense to consider SUTVA as violated within a cluster: there is no way for one candidate in a district to be treated by a different district-party ideology than other candidates.¹⁰ The positioning of one candidate may also affect the positioning of another, which could violate the “no interference” component of SUTVA. Under this violation, the treatment effect at the individual level is not identified. If SUTVA holds *between* groups, however, it is possible to identify a treatment effect by considering average effects across groups (Hill n.d.). In potential outcomes notation, even if we can define potential outcomes at the individual level ($CF_i(\bar{\theta}_{g[i]})$), the lowest level where we could credibly *identify* treatment effects would be the group level, where the potential outcome for a group is the average outcome within the group ($\overline{CF}_g(\bar{\theta}_g)$). This is consistent with the multilevel model setup that we have so far, where the ACDE is a function of aggregate data and aggregate parameters only (see Equation (4.10)).

There are a few additional considerations for causal inference with hierarchical data that, although I do not pursue these threads in this project, could be relevant for future work with similar data. A correlation between treatment effects and group size may arise if a crowded primary field causes larger treatment effects because stiffer competition leads candidates to be more responsive to district-party ideology. On the other hand, more crowded fields would lead to smaller treatment effects if candidates take heterogeneous ideological positions to differentiate themselves. If treatment effects are correlated with group size, then the average causal effect for a candidate is not equivalent to average difference among groups. Instead, the average effect for candidates must be a size-weighted average of group effects (Hill n.d.). I do not pursue this possibility in this project because these dynamics are not identifiable with data on primary candidates only, since incumbents may take ideological positions to deter challengers even if no challengers actually emerge. As such, the observed number of candidates in a district may not capture the true degree of primary threat (Hirano et al. 2010;

¹⁰Candidates may vary in their ability to perceive district-party ideology, but that might also be described as an issue of treatment compliance or treatment effect heterogeneity.

Maisel and Stone 1997; Stone and Maisel 2003).

One additional consideration for group-level effects is the possibility that group size affects treatment assignment. This may be true if the long-run dynamics of primary competition within a district-party have feedback effects on local ideology, for instance if partisan constituents become more ideologically aware by experiencing stronger intra-party competition in their district, or less ideological after a long period of representation by a single incumbent with little primary competition. There is evidence that primaries contain more ideological campaign content in certain periods of heightened partisan mobilization (Boatright 2013), which could increase voters' ideological awareness as well. Whether voters are responding to primary competition in the district *per se* or to a national state of partisan agitation is an interesting but thoroughly challenging question for future researchers to explore, were they to extend the data and methods in this project to a greater number of election cycles and dynamic causal modeling approaches (e.g. Blackwell and Glynn 2018; Imai and Kim 2019).

4.2.4 Data

For the CF score outcome measure, I specifically use the dynamic CF score provided by Bonica (2019b), which is re-estimated for each two-year FEC cycle. The measure of district-level partisanship, the mediator in the sequential *g*-estimation routine, is the two-party Republican presidential vote share from the previous election cycle, which is provided and matched to candidacies by Bonica (2019b). District-party ideology $\tilde{\theta}_g$, the treatment variable, is measured from the ideal point model in Chapter 2.

Pre-treatment confounders \mathbf{x}_d are included to identify the effect of the presidential vote in stage 1 and the effect of district-party ideology in stage 2. These covariates were organized and matched to primary candidacies by the Primary Timing Project by Boatright, Moscardelli, and Vickrey (2017). District demographic variables (sourced originally from the American Community Survey) include district median income, population density, and land area,

as well the percent of a district population that is White, Latino, college educated, below the federal poverty line, unemployed, employed in the service industry, employed in blue-collar jobs, aged 28–24, and aged 65+. The Boatright, Moscardelli, and Vickrey (2017) data also provide Mayhew (1986)’s five-level “traditional party organization” and binary “persistent factionalism” classifications for state level (Mayhew 1986). The stage-1 model also contains intermediate controls \mathbf{z}_d for identifying the effect of the past presidential vote on CF scores. These variables include the district-party ideal point for the *other* party group in the same district, because both parties should partially affect aggregate district voting, and year fixed effects to capture average shifts in CF scores that are correlated with average shifts in presidential voting.¹¹

All controls except for the fixed effects and binary indicators from Mayhew (1986) are centered at their means and scaled by their standard deviations. This makes it easier to specify priors for coefficients, since standardized coefficients are unlikely to exceed a 1 except when predictors are highly correlated. District-party ideology $\bar{\theta}_g$, the treatment variable, is measured on its original scale anchored by the item parameters in the Chapter 2 model. The presidential vote variable is centered at its reference value for demediation, 0.5, and then divided by 10 so that a one-unit change in the model represents a 10 point change in vote share. CF scores are measured on their original scale, which spans roughly as wide as -5 to 5 across both parties, both the vast majority of Democrats occupy values in $[-2, 0]$ and Republicans in $[0, 2]$.

4.3 Findings

I estimate models in several subsets of data. First, I estimate separate models for all Democratic and all Republican candidates in the sample. I then estimate models that divide each

¹¹Fixed effects are not included in pre-treatment controls because district-party ideology is fixed across the redistricting cycle, so they do not improve causal identification in stage 2.

Table 4.1: Sample sizes in all estimated models.

Full Sample	By Incumbency Status	By Primary Rules
Democrats = 1,970	Incumbents (D = 568, R = 704)	Closed (D = 933, R = 1,026)
Republicans = 2,192	Challengers (D = 1,089, R = 1,096)	Semi-Closed (D = 332, R = 375)
	Open Seats (D = 313, R = 392)	Open (D = 705, R = 791)

partisan subsample into incumbent candidates, challengers of incumbents, and candidates running for an open congressional seat. Finally, I estimate separate models for candidates running in closed primaries, semi-closed primaries, and open primaries. Table 4.1 shows the sample sizes in each model subset.

For the sake of computation time, I estimate these models by approximating the posterior distribution using the mean-field variational Bayes routine available in Stan (Kucukelbir et al. 2015). Variational Bayesian inference (VB) finds and optimizes a simpler distribution that is similar to the true posterior distribution in terms of Kullback–Leibler divergence (Grimmer 2011). Variational estimators are asymptotically consistent and can be used to estimate any Stan model. But because they require approximations, samples from the approximate distribution tend to underestimate the variance in the true posterior distribution (Wang and Titterton 2012). The benefit, however, is that models that would take hours to estimate using MCMC can be estimated using VB in roughly one minute, which is extremely valuable for building and estimating the 14 models included in this chapter.

Before turning to the results, recall that a distinguishing feature of the Bayesian approach is the use of priors to represent measurement error in district-party ideal points. This is an intuitive solution to the problem of uncertainty propagation because uncertainty in causal effects naturally reflects uncertainty in the data by marginalizing over the ideal points. Another interesting consequence is that the posterior distribution for the ideal points could be different from the prior, depending on the information that the inferential the model can provide about the values of the ideal point parameters. I plot the prior and posterior

distributions for ideal points alongside one another in Figure 4.9, using points to represent prior and posterior means and bars to represent uncertainty intervals. Posterior estimates come from the models estimated on the full samples of Democrats and Republicans. The data fall along the 45-degree line, indicating that prior and posterior ideal points are similar to one another, which is a sensible result that increases our confidence in the computational accuracy of the model. The similarity between prior and posterior ideal points is also convenient for understanding the regression results, because coefficients can be interpreted in the original scale of the data without any need to post-process results into a familiar scale.¹²

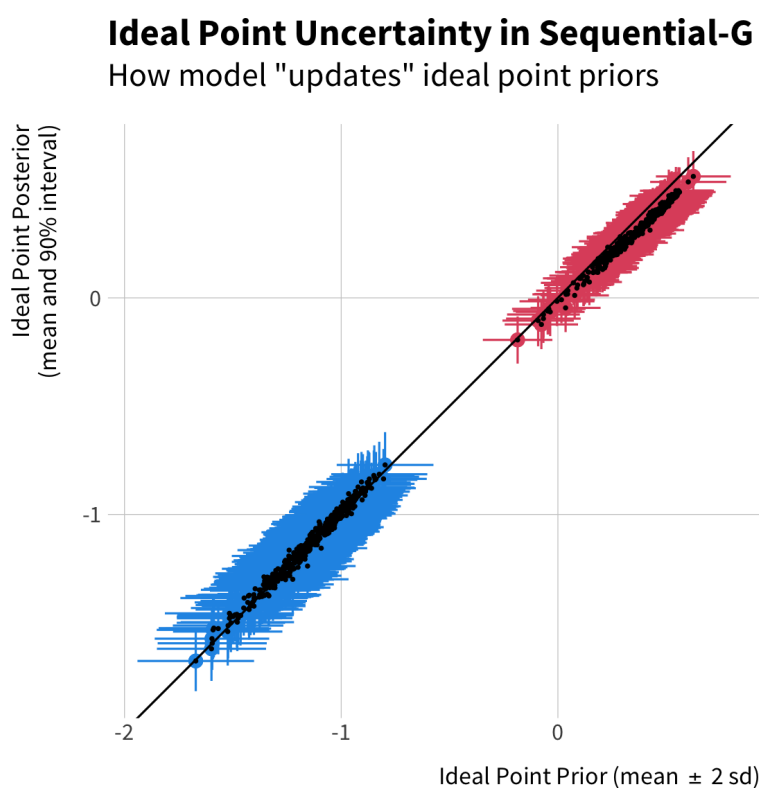


Figure 4.9: Marginal prior and posterior distributions for ideal points in sequential g model.

¹²In a regression context like this, the ideal points and their regression coefficient are not mutually identifiable because each could be arbitrarily scaled in offsetting ways. Estimation with MCMC may suffer under this non-identifiability by “wandering” through the weakly identified regions of parameter space, which can be corrected by post-processing parameter samples by applying scale constraints to each MCMC interaction. The variational algorithm is deterministic and does not wander in the same way, which eliminates the need to post-process estimates to recover sensible answers.

We now turn to the sequential g results. Figure 4.10 plots VB posterior means and 90% intervals for the full sample of Democratic candidates. Stage 1 parameters are plotted as squares, and stage 2 parameters are plotted as circles. The top-left panel contains the key parameters most relevant to causal inference, including the coefficients for the mediator and treatment variables from both stages of the estimation. The bottom-left panel plots the standard deviations of the district errors and residual errors. The right-side panel contains regression coefficients for all control variables.

Sequential G Parameters

All Democratic Candidates

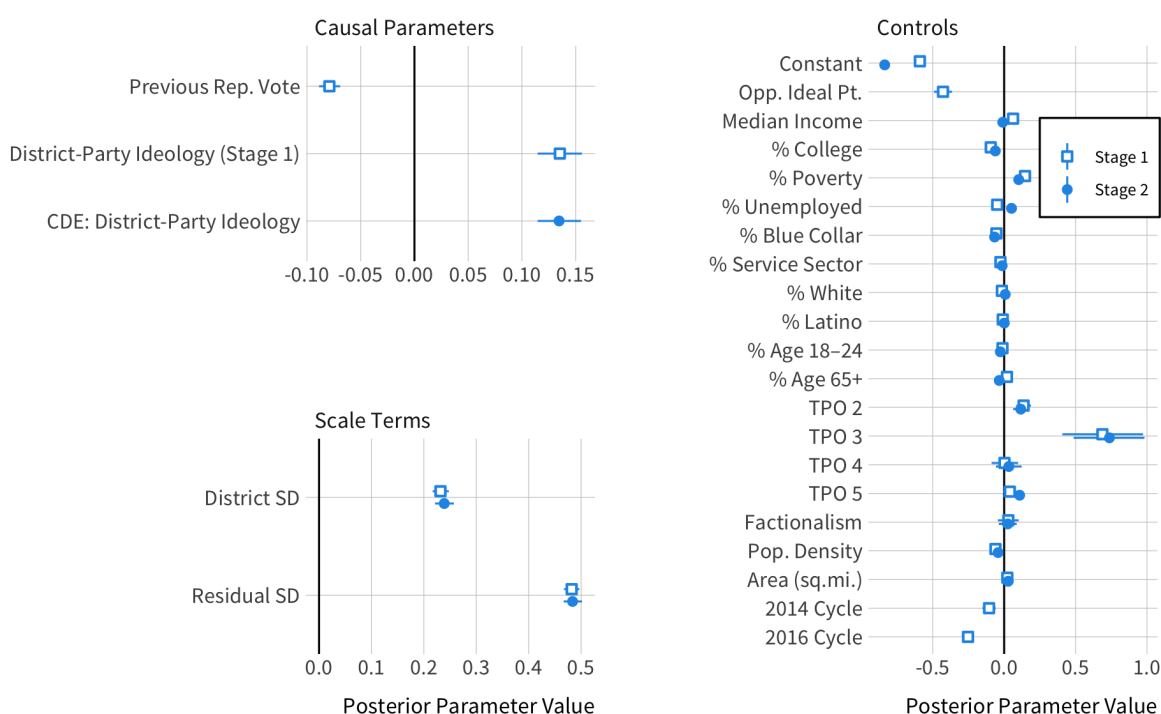


Figure 4.10: Sequential g results for Democratic candidates. Points and intervals are variational point estimates and 90 percent intervals from the approximate posterior distribution.

Under sequential ignorability and no intermediate interactions assumptions, stage 1 estimates the effect of the past presidential vote share on CF scores. Interestingly, this coefficient

is negative, indicating that Democratic primary candidates running in districts that vote more heavily Republican are actually more *progressive* than Democratic candidates who represent more Democratic-leaning districts. This is an unexpected relationship given the general findings in the literature that candidates take more moderate stances in more moderate districts, which should manifest as a positive coefficient. It is important to keep in mind that this sample of data contains incumbent and non-incumbent candidates, as well as primary winners and losers. Because most primary elections do not occur under highly competitive circumstances, the typical primary campaign may not resemble the predictions from “Downsian” formal models that assume perfect candidate competition (Burden 2004). The mixture of incumbent and non-incumbent candidates may weight the sample more heavily toward candidates who corner more extreme or idiosyncratic candidate positions in their attempts to attract attention away from incumbents, who are likely to position themselves more in line with overall district preferences (Ansolabehere, Snyder, and Stewart 2001). The models that stratify on incumbency in Section 4.3.1 are consistent with these possibilities. One other possibility is that the introduction of my new measure, district-party ideology, is responsible. If candidates are indeed sensitive to *partisan* preferences in their district, and partisan preferences are positively correlated with district voting on average, then the strong relationship between candidate positions and the district vote was at least partially confounded in all past studies. The results in Figure 4.10 are consistent with that possibility, since the coefficient on district-party ideology is large and positive even in stage 1, when it does not have a causal interpretation but is instead included as a control to identify the presidential vote effect.

The controlled direct effect of district-party ideology is positive, with a posterior mean indicating that one-unit increase in district-party ideology causing a 0.13-unit increase in CFscores. Using standardized coefficients, the posterior mean suggests that an increase of one standard deviation in district-party ideology causes 0.04-standard deviation increase in

CF scores, which is a substantively small effect.

The control coefficients do not have causal interpretation, and even interpreting them as partial correlations is problematic because of collider bias. Nonetheless readers might find some of the coefficient estimates intriguing or intuitive given the makeup of the Democratic Party coalition and the polarized environment of the 2010s. The opposing party ideal point mean has a clear negative coefficient, indicating that more conservative Republicans coincide on average with more progressive Democrats. Progressivism among Democrats is greater in districts with greater density, less land area, and greater numbers of service sector employees. The year fixed effects suggest candidates are more liberal in more recent years, which fits a pattern of polarization over time. Some interesting or counter-intuitive findings are that Democrats are more progressive in districts with more college graduates and more blue-collar workers, and they are more conservative in districts with higher poverty.

Turning to the Republican estimates in Figure 4.11, we find a similar pattern among the causal parameters. The presidential vote is again inversely related to ideal points, with greater Republican voting causing more conservative Republican candidates under the causal assumptions. We also find a positive controlled direct effect of district-party ideology in the stage 2 model, indicating that Republicans run more conservative campaigns in more conservative districts. The district-party ideology effect in the Republican Party is larger than the effects in the Democratic Party on the scale of the raw data, with a coefficient of 0.27 (versus 0.13). In standardized coefficients, the effects are more similar, with a Republican coefficient of 0.06 versus the Democratic coefficient of 0.04. This is because district-party ideology is more similar across Republican groups than Democratic groups, with a standard deviation in ideal point means of 0.11 versus 0.17, meaning that one standardized unit increase is a smaller increase in absolute terms among Republicans than among Democrats.

Benchmarking these cross-party comparisons using raw or standardized coefficients has consequences for the conclusions we can draw about representation in the two parties. One

Sequential G Parameters

All Republican Candidates

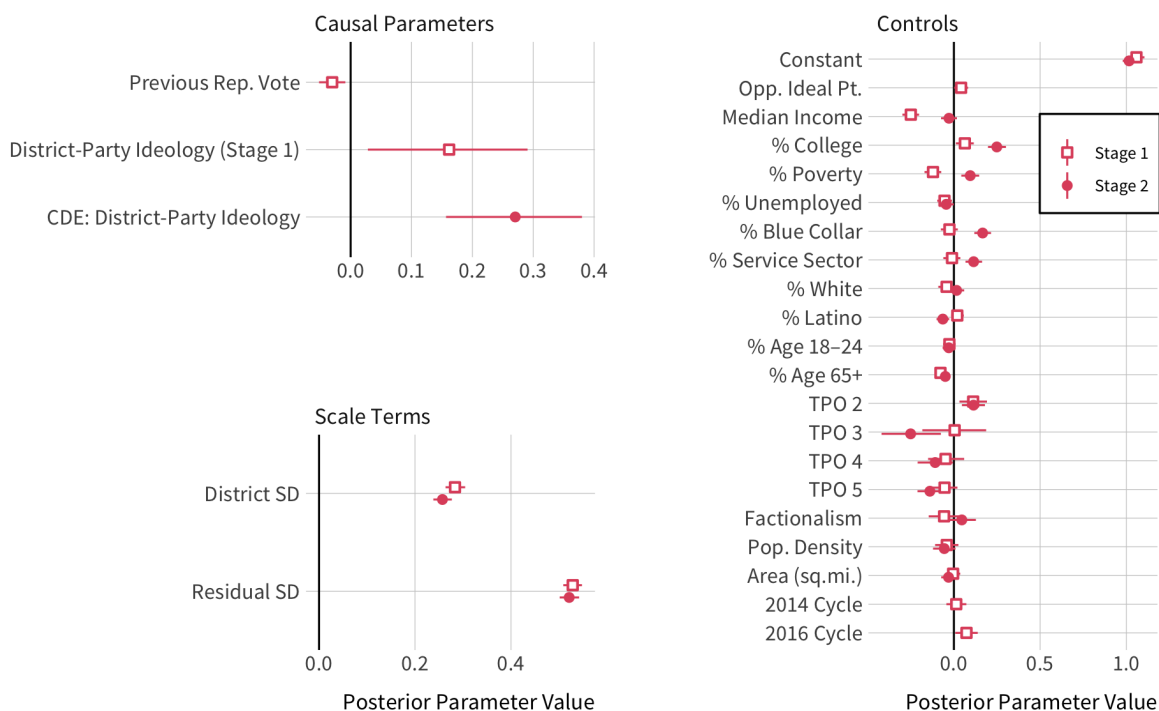


Figure 4.11: Sequential g results for Republican candidates. Points and intervals are variational point estimates and 90 percent intervals from the approximate posterior distribution.

interpretation places greater emphasis on the standardized picture: the initial appearance of greater responsiveness among Republicans using raw coefficients (e.g. Clinton 2006) makes an incorrect assumption that the scope of ideological conflict is the same among Democrats and Republicans, when in fact Republican voters and elites are much more ideologically cohesive than Democrats (Lelkes and Sniderman 2016). But conditioning on the actual scope of ideological conflict in the two parties, Republican and Democratic candidates respond to their constituencies with proportional adjustments to their campaign positions. Another interpretation places greater emphasis on the raw coefficients: if the scope of ideological conflict within the Democratic Party is greater, this is essential to keep in mind

for understanding how candidates position themselves in relation to voters. If the Democratic Party contains a larger variety of conflicting group interests, and elite party members have to reign these interests into a coherent platform, the result could be an elite party that appears to be insensitive to the big-tent heterogeneity in voters' policy preferences.

4.3.1 Effect modification: incumbency status and primary rules

This section examines the results for models estimated within strata defined by candidate incumbency status and state primary rules. It is an important to note that conditional effects do not necessarily represent the causal effects of stratum membership, a distinction that is often overlooked in examinations of heterogeneous causal effects (Kam and Trussler 2017). Differences in the CDE across incumbency or primary rules reflect causal heterogeneity, but the sources of heterogeneity could come from factors that confound the the link between strata and outcomes. As such, these results should be seen from an “effect modification” point of view. In order to claim that incumbency or primary rules *cause* heterogeneity in the effects of district-party ideology requires additional assumptions that incumbency or primary rules are ignorable, which are tasks large enough to fill separate dissertations altogether.

I first examine effects by stratifying across candidate status, estimating separate models for incumbents, challengers of incumbents, and open seat candidates. Prevailing literature that examines incumbency and candidate positioning have diverging predictions for these results. Ansolabehere, Snyder, and Stewart (2001) find that incumbents take more moderate campaign positions than non-incumbent candidates, with challengers taking more extreme than both incumbents and open seat candidates. This reflects a perspective that incumbency selects for candidates that take median positions that ensure their reelection, while challengers must take more innovative and extreme stances to garner attention. Burden (2004) finds instead that incumbents take more partisan positions, perhaps because incumbency provides other advantages that give incumbents more security to take positions that don't match district

preferences as closely.

The findings in this study contain a mixture of evidence for against each of these views. Figure 4.12 plots the mediator effect and controlled direct effect for incumbents, challengers, and open seat candidates. Among Democrats, incumbents appear to follow patterns of positioning dictated by the SPD. Coefficients for the presidential vote and district-party ideology are both positive, indicating that Democrats run as more progressive candidates in districts with higher Democratic presidential voting and with more progressive partisan constituents. Challenger positioning is inversely related to the presidential vote but positively related to district-party ideology, indicating that challengers target the partisan constituency at the expense of the general election constituency. This is consistent with an account of incumbent–challenger dynamics where challengers take extreme positions to garner attention to themselves against a more moderate, or that the responsiveness of incumbents results from the selection of well-fitting candidates by district voting (Ansolabehere, Snyder, and Stewart 2001). Open seat candidates behave more similarly to incumbents than to challengers, which contradicts the finding by Ansolabehere, Snyder, and Stewart (2001) that open seat candidates may be most “out of step” with district preferences, and may be more consistent with the Burden (2004) argument that non-incumbents position themselves most aggressively to match voter preferences because they have fewer build-in advantages to overcome any ideological mis-fits.

Among Republicans, the results suggest that incumbent candidates are responsive to within-party preferences, but essentially unrelated to aggregate district voting. This is consistent with a Clinton (2006) view that Republicans position themselves to fit their partisan constituencies, and correlation with district voting is incidental. This also fits with the Burden argument that incumbency provides insurance that lets incumbents deviate from district-optimal preferences, which is also consistent with the finding that Republican challengers are more responsive to district voting than Republican incumbents. Results among Republican

Effect Modification by Incumbency Status

Mediator Effects and Controlled Direct Effects

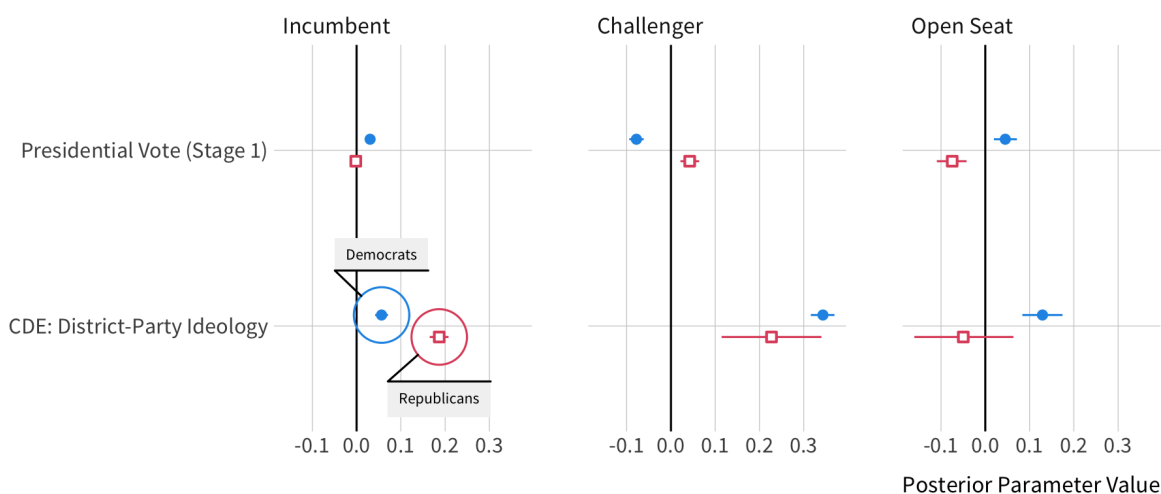


Figure 4.12: Sequential g results for incumbents, challengers, and open seat candidates. Points and intervals are variational point estimates and 90 percent intervals from the approximate posterior distribution.

incumbents are difficult to interpret, since the results show a negative effect of Republican voting and no clear relationship between district-party conservatism and candidate conservatism. Overall, these results suggest that nearly candidates of all incumbency status are responsive on average to district-party ideology, which supports an SPD view of primary competition. Among incumbents, the greater degree of general-election responsiveness by Democrats than Republicans tracks Clinton (2006), but the non-uniform findings among challengers and incumbents lack a clear interpretation.

We now turn to an examination of primary rules. Past studies of primary rules typically find no clear effects of different rules either on candidate positioning or vote choice. These studies never measure district-party ideology, however, so has never been clear whether candidates are in fact more responsive to district-party ideology in less-open primary systems. My results in Figure 4.13 suggest that the inclusion of district-party preferences is consequential for understanding primary openness. Among Democrats, the controlled direct effect

of district party ideology is strongest in closed systems and weakest in open systems, with semi-closed systems in the middle. This monotonic relationship is exactly consistent with the conventional wisdom on primary openness and candidate positioning. The results among Republicans are not monotonic—the strongest relationship is among semi-closed rather than closed systems—but generally support a conclusion that candidates in open systems are less responsive to district-party ideology, where the relationship between district-parties and candidates is actually negative.

Effect Modification by Primary Openness

Controlled Direct Effects of District-Party Ideology

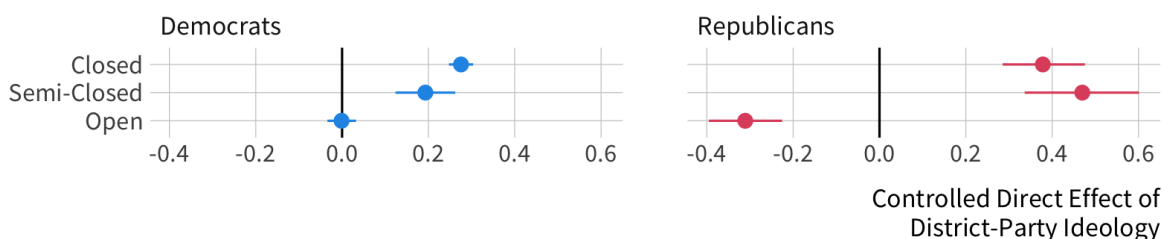


Figure 4.13: Sequential g results for candidates in states with closed, semi-closed, and open primary systems. Points and intervals are variational point estimates and 90 percent intervals from the approximate posterior distribution.

It is worth reiterating that these patterns across primary rules should not be interpreted as causal, although it is notable that they diverge from the majority of recent research on the effects of primary rules. Furthermore, even though candidates may position themselves in accordance with the strategic environment created by primary institutions, this is no guarantee that primary voters recognize which candidate best fits their ideological preferences conditional on a slate of primary candidates. This is a question I revisit in Chapter 5.

4.4 Discussion

This study contained three notable contributions for the study of the strategic positioning dilemma and primary candidate positioning. First, I employ novel measures of district-party ideology, a concept that is essential to the SPD theory but never operationalized in existing studies to date. Second, I advance the causal credibility of these studies by positing a causal estimand and implementing a statistical approach that estimates it under explicitly stated assumptions. And third, I demonstrate how to bring this method into a Bayesian causal framework, incorporating measurement error in the key independent variable as a prior distribution and deriving a posterior distribution that summarizes uncertainty in all model parameters. I find that primary candidates do position themselves to fit district-party policy preferences, even when controlling for aggregate district voting as an intermediary causal mechanisms. Furthermore, I find that candidate responsiveness is generally greater in states with stricter primary participation rules, a hypothesis that is common among political punditry but doubted among primary elections research.

The contributions of this study provide a structure for improving the research further. Causal diagrams are valuable for understanding how theories of primaries other than the SPD could manifest in the data. Consider, for instance, the notion that candidates position themselves to earn the good favor and resources of policy-demanding groups in the party network. Whether this theoretical perspective jeopardize the causal interpretation of the preceding analysis depends on where policy-demanding are located in a causal diagram. If intermediary groups are an outgrowth of the social, economic, and other demographic bases of the district—so they are descendants of variables x_d —and partisan voters take opinion cues from these groups' activities in the district, then failing to control for intermediary group influence may lead to bias in the estimated effects of district-party ideology.

This study could also be improved by collecting more years of data and setting up a

different research design based on within-unit variation, such as a difference-in-differences panel design. Unfortunately, most routine implementations of difference-in-differences modeling are problematic under data with variation in treatment timing, an issue that econometricians are only just starting to understand (Goodman-Bacon 2018). Panel data approaches to causal inference in political science either generalize the “demediation” routine already employed in this analysis (Blackwell and Glynn 2018) or use fixed-effects models that require particular assumptions how past treatments and outcomes are allowed to affect future treatments and future outcomes (Imai and Kim 2019). Modernizing the primary representation literature for a new generation of causal approaches will not be a simple, one-off endeavor.

Even if researchers stick to the single-stage demediation approach, they could relax modeling assumptions by moving away from the linear modeling laid out by Acharya, Blackwell, and Sen (2016) and implemented in this chapter. Because the identification assumptions, demediation function, and ACDE are nonparametrically defined in terms of expectations about data, the method could be implemented using more flexible estimators of conditional expectations such as matching or machine learning methods.

— 5 —

District-Party Ideology and Primary Election Outcomes

References

- Abramowitz, Alan I, and Kyle L Saunders. 1998. "Ideological realignment in the us electorate." *The Journal of Politics* 60(03): 634–652.
- Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "Explaining causal findings without bias: Detecting and assessing direct effects." *American Political Science Review* 110(3): 512–529.
- Ahler, Douglas J, Jack Citrin, and Gabriel S Lenz. 2016. "Do open primaries improve representation? An experimental test of california's 2012 top-two primary." *Legislative Studies Quarterly* 41(2): 237–268.
- Aldrich, John H. 1983. "A downsian spatial model with party activism." *American Political Science Review* 77(04): 974–990.
- Aldrich, John H. 2011. *Why parties?: A second look*. University of Chicago Press.
- American Political Science Association, Committee on Political Parties. 1950. *Toward a more responsible two-party system*. Johnson Reprint Company.
- Ansolabehere, Stephen et al. 2010. "More democracy: The direct primary and competition in us elections." *Studies in American Political Development* 24(02): 190–205.

- Ansolahehere, Stephen, Jonathan Rodden, and James M Jr Snyder. 2008. "The strength of issues: Using multiple measures to gauge preference stability, ideological constraint, and issue voting." *American Political Science Review*: 215–232.
- Ansolahehere, Stephen, James M Snyder, and Charles Stewart. 2001. "Candidate positioning in U.S. house elections." *American Journal of Political Science*: 136–159.
- Athey, Susan, Guido W Imbens, and Stefan Wager. 2016. "Approximate residual balancing: De-biased inference of average treatment effects in high dimensions." *arXiv preprint arXiv:1604.07125*.
- Barber, Michael J. 2016. "Ideological donors, contribution limits, and the polarization of american legislatures." *The Journal of Politics* 78(1): 296–310.
- Barber, Michael J, Brandice Canes-Wrone, and Sharece Thrower. 2016. "Ideologically sophisticated donors: Which candidates do individual contributors finance?" *American Journal of Political Science*.
- Barber, Michael, and Jeremy C Pope. 2019. "Does party trump ideology? Disentangling party and ideology in america." *American Political Science Review*: 1–17.
- Barberá, Pablo. 2015. "Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data." *Political analysis* 23(1): 76–91.
- Bartels, Larry M. 2000. "Partisanship and voting behavior, 1952-1996." *American Journal of Political Science*: 35–50.
- Bartels, Larry M. 2009. *Unequal democracy: The political economy of the new gilded age*. Princeton University Press.

- Beck, Nathaniel, Gary King, and Langche Zeng. 2004. "Theory and evidence in international conflict: A response to de marchi, gelpi, and grynaviski." *American Political Science Review*: 379–389.
- Betancourt, Michael. 2017. "A conceptual introduction to hamiltonian monte carlo." *arXiv preprint arXiv:1701.02434*.
- Betancourt, Michael. 2019. "The convergence of markov chain monte carlo methods: From the metropolis method to hamiltonian monte carlo." *Annalen der Physik* 531(3): 1700214.
- Betancourt, Michael. 2018. "Towards a principled bayesian workflow."
- Betancourt, Michael, and Mark Girolami. 2015. "Hamiltonian monte carlo for hierarchical models." *Current trends in Bayesian methodology with applications* 79: 30.
- Black, Duncan. 1948. "On the rationale of group decision-making." *The Journal of Political Economy*: 23–34.
- Blackwell, Matthew, and Adam N Glynn. 2018. "How to make causal inferences with time-series cross-sectional data under selection on observables." *American Political Science Review* 112(4): 1067–1082.
- Blackwell, Matthew, James Honaker, and Gary King. 2017. "A unified approach to measurement error and missing data: Overview and applications." *Sociological Methods & Research* 46(3): 303–341.
- Boatright, Robert G. 2013. *Getting primaried: The changing politics of congressional primary challenges*. University of Michigan Press.
- Boatright, Robert G, Vincent G Moscardelli, and Clifford Vickrey. 2017. "The consequences of primary election timing." *Primary Timing Project*, June.

- Bonica, Adam. 2019a. "Are donation-based measures of ideology valid predictors of individual-level policy preferences?" *The Journal of Politics* 81(1): 327–333.
- Bonica, Adam. 2019b. "Database on ideology, money in politics, and elections: Public version 1.0."
- Bonica, Adam. 2013. "Ideology and interests in the political marketplace." *American Journal of Political Science* 57(2): 294–311.
- Bonica, Adam. 2014. "Mapping the ideological marketplace." *American Journal of Political Science* 58(2): 367–386.
- Bonica, Adam. 2020. "Why are there so many lawyers in congress?" *Legislative Studies Quarterly* 45(2): 253–289.
- Brady, David W, Hahrie Han, and Jeremy C Pope. 2007. "Primary elections and candidate ideology: Out of step with the primary electorate?" *Legislative Studies Quarterly* 32(1): 79–105.
- Broockman, David E. 2016. "Approaches to studying policy representation." *Legislative Studies Quarterly* 41(1): 181–215.
- Broockman, David E, and Christopher Skovron. 2018. "Bias in perceptions of public opinion among political elites." *American Political Science Review* 112(3): 542–563.
- Brunell, Thomas L. 2006. "Rethinking redistricting: How drawing uncompetitive districts eliminates gerrymanders, enhances representation, and improves attitudes toward congress." *PS: Political Science and Politics* 39(1): 77–85.
- Brunell, Thomas L, Bernard Grofman, and Samuel Merrill. 2016. "Components of party polarization in the us house of representatives." *Journal of Theoretical Politics* 28(4): 598–624.

- Bullock, Will, and Joshua D Clinton. 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.
- Burden, Barry C. 2004. "Candidate positioning in u.s. Congressional elections." *British Journal of Political Science* 34(02): 211–227.
- Burden, Barry C. 2001. "The polarizing effects of congressional primaries." *Congressional Primaries and the Politics of Representation*: 95–115.
- Burden, Barry C, Gregory A Caldeira, and Tim Groseclose. 2000. "Measuring the ideologies of us senators: The song remains the same." *Legislative Studies Quarterly*: 237–258.
- Butler, Daniel M, and Eleanor Neff Powell. 2014. "Understanding the party brand: Experimental evidence on the role of valence." *The Journal of Politics* 76(2): 492–505.
- Bürkner, Paul-Christian, and others. 2017. "Brms: An r package for bayesian multilevel models using stan." *Journal of Statistical Software* 80(1): 1–28.
- Campbell, Angus et al. 1960. New York: John Wiley and Sons 77 *The american voter*.
- Canes-Wrone, Brandice, David W Brady, and John F Cogan. 2002. "Out of step, out of office: Electoral accountability and house members' voting." *American Political Science Review* 96(01): 127–140.
- Canes-Wrone, Brandice, William Minozzi, and Jessica Bonney Reveley. 2011. "Issue accountability and the mass public." *Legislative Studies Quarterly* 36(1): 5–35.
- Carlson, David. 2020. "Estimating a counter-factual with uncertainty through gaussian process projection."

- Carpenter, Bob et al. 2016. "Stan: A probabilistic programming language." *Journal of Statistical Software* 20: 1–37.
- Caughey, Devin, James Dunham, and Christopher Warshaw. 2018. "The ideological nationalization of partisan subconstituencies in the american states." *Public Choice* 176(1-2): 133–151.
- Caughey, Devin, and Christopher Warshaw. 2015. "Dynamic estimation of latent opinion using a hierarchical group-level irt model." *Political Analysis* 23(2): 197–211.
- Caughey, Devin, and Christopher Warshaw. 2018. "Policy preferences and policy change: Dynamic responsiveness in the american states, 1936–2014."
- Caughey, Devin, and Christopher Warshaw. 2019. "Public opinion in subnational politics." *The Journal of Politics* 81(1): 352–363.
- Chernozhukov, Victor et al. 2017. "Double/debiased/neyman machine learning of treatment effects." *American Economic Review* 107(5): 261–65.
- Clinton, Joshua D. 2006. "Representation in congress: Constituents and roll calls in the 106th house." *Journal of Politics* 68(2): 397–409.
- Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. "The statistical analysis of roll call data." *American Political Science Review* 98(02): 355–370.
- Cohen, Marty et al. 2009. *The party decides: Presidential nominations before and after reform*. University of Chicago Press.
- Cox, Gary W. 1990. "Centripetal and centrifugal incentives in electoral systems." *American Journal of Political Science*: 903–935.

- Cox, Gary W, and Mathew D McCubbins. 2005. *Setting the agenda: Responsible party government in the us house of representatives*. Cambridge University Press.
- Doherty, David, Conor M Dowling, and Michael G Miller. 2019. "Do local party chairs think women and minority candidates can win? Evidence from a conjoint experiment." *The Journal of Politics* 81(4): 1282–1297.
- Downs, Anthony. 1957. *An economic theory of democracy*. New York: Harper; Row.
- Duane, Simon et al. 1987. "Hybrid monte carlo." *Physics letters B* 195(2): 216–222.
- Ellis, Christopher, and James A Stimson. 2012. *Ideology in america*. Cambridge University Press.
- Enns, Peter K, and Julianna Koch. 2013. "Public opinion in the us states: 1956 to 2010." *State Politics & Policy Quarterly* 13(3): 349–372.
- Epstein, Lee et al. 2007. "The judicial common space." *Journal of Law, Economics, and Organization* 23(2): 303–325.
- Feldman, Stanley, and John Zaller. 1992. "The political culture of ambivalence: Ideological responses to the welfare state." *American Journal of Political Science*: 268–307.
- Fenno, Richard F. 1978. *Home style: House members in their districts*. Pearson College Division.
- Fiorina, Morris P, Samuel J Abrams, and Jeremy C Pope. 2005a. *Culture war? The myth of a polarized america*. Pearson Longman New York.
- Fiorina, Morris P, Samuel J Abrams, and Jeremy C Pope. 2005b. *Culture war? The myth of a polarized america*. Pearson Longman New York.
- Fowler, Anthony, and Andrew B Hall. 2016. "The elusive quest for convergence." *Quarterly Journal of Political Science* 11: 131–149.

- Fowler, Linda L. 1982. "How interest groups select issues for rating voting records of members of the us congress." *Legislative Studies Quarterly*: 401–413.
- Free, Lloyd A, and Hadley Cantril. 1967. "The political beliefs of americans."
- Freeman, Jo. 1986. "The political culture of the democratic and republican parties." *Political Science Quarterly* 101(3): 327–356.
- Gabry, Jonah et al. 2019. "Visualization in bayesian workflow." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 182(2): 389–402.
- Geer, John G. 1988. "Assessing the representativeness of electorates in presidential primaries." *American Journal of Political Science*: 929–945.
- Gelman, Andrew, and Jennifer Hill. 2006. *Data analysis using regression and multi-level/hierarchical models*. Cambridge University Press.
- Gelman, Andrew, Daniel Simpson, and Michael Betancourt. 2017. "The prior can often only be understood in the context of the likelihood." *Entropy* 19(10): 555.
- Gerring, John. 2001. *Social science methodology: A criterial framework*. Cambridge University Press.
- Gilens, Martin, and Benjamin I. Page. 2014. "Testing theories of american politics: Elites, interest groups, and average citizens." *Perspectives on Politics* 12(3).
- Gill, Jeff. 2014. *20 Bayesian methods: A social and behavioral sciences approach*. CRC press.
- Goodman-Bacon, Andrew. 2018. *Difference-in-differences with variation in treatment timing*. National Bureau of Economic Research.
- Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan hearts and minds*. New Haven, CT: Yale University Press.

- Green, Donald P et al. 2016. "The effects of lawn signs on vote outcomes: Results from four randomized field experiments." *Electoral Studies* 41: 143–150.
- Greenland, Sander, Judea Pearl, and James M Robins. 1999. "Causal diagrams for epidemiologic research." *Epidemiology*: 37–48.
- Grimmer, Justin. 2011. "An introduction to bayesian inference via variational approximations." *Political Analysis* 19(1): 32–47.
- Grossman, Matthew, and David A. Hopkins. 2016. Oxford University Press *Asymmetric politics: Ideological republicans and group interest democrats*.
- Grosz, Michael P, Julia M Rohrer, and Felix Thoemmes. 2020. "The taboo against explicit causal inference in nonexperimental psychology."
- Hacker, Jacob S, Paul Pierson, and others. 2005. *Off center: The republican revolution and the erosion of american democracy*. Yale University Press.
- Hahn, P Richard et al. 2018. "Regularization and confounding in linear regression for treatment effect estimation." *Bayesian Analysis* 13(1): 163–182.
- Hahn, P Richard et al. 2020. "Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects." *Bayesian Analysis*.
- Hall, Andrew B. 2015. "What happens when extremists win primaries?" *American Political Science Review* 109(01): 18–42.
- Hall, Andrew B, and James M Snyder. 2015. "Candidate ideology and electoral success. Working paper: https://dl.dropboxusercontent.com/u/11481940/hall_snyder_ideology.pdf."

- Hall, Andrew B, and Daniel M Thompson. 2018. "Who punishes extremist nominees? Candidate ideology and turning out the base in us elections." *American Political Science Review* 112(3): 509–524.
- Henderson, John A. 2016. "An experimental approach to measuring ideological positions in political text." *Available at SSRN 2852784*.
- Hernán, Miguel A. 2018. "The c-word: Scientific euphemisms do not improve causal inference from observational data." *American journal of public health* 108(5): 616–619.
- Hill, Jennifer. "Multilevel models and causal inference." In *The SAGE handbook of multilevel modeling*, SAGE Publications Ltd, p. 201–220. <https://doi.org/10.4135/9781446247600.n12>.
- Hill, Jennifer L. 2011. "Bayesian nonparametric modeling for causal inference." *Journal of Computational and Graphical Statistics* 20(1): 217–240.
- Hill, Seth J. 2015. "Institution of nomination and the policy ideology of primary electorates." *Quarterly Journal of Political Science* 10(4): 461–487.
- Hill, Seth J, and Gregory A Huber. 2017. "Representativeness and motivations of the contemporary donorate: Results from merged survey and administrative records." *Political Behavior* 39(1): 3–29.
- Hirano, Shigeo et al. 2010. "Primary elections and party polarization." *Quarterly Journal of Political Science* 5: 169–191.
- Hirano, Shigeo, and Michael M Ting. 2015. "Direct and indirect representation." *British Journal of Political Science* 45(3): 609.
- Horiuchi, Yusaku, Kosuke Imai, and Naoko Taniguchi. 2007. "Designing and analyzing randomized experiments: Application to a japanese election survey experiment." *American Journal of Political Science* 51(3): 669–687.

- Imai, Kosuke, and In Song Kim. 2019. "When should we use unit fixed effects regression models for causal inference with longitudinal data?" *American Journal of Political Science* 63(2): 467–490.
- Jackman, Simon. 2009. 846 *Bayesian analysis for the social sciences*. John Wiley & Sons.
- Jacobson, Gary C. 2012. "The electoral origins of polarized politics: Evidence from the 2010 cooperative congressional election study." *American Behavioral Scientist* 56(12): 1612–1630.
- Kam, Cindy D, and Marc J Trussler. 2017. "At the nexus of observational and experimental research: Theory, specification, and analysis of experiments with heterogeneous treatment effects." *Political Behavior* 39(4): 789–815.
- Kastellec, Jonathan P et al. 2015. "Polarizing the electoral connection: Partisan representation in supreme court confirmation politics." *The journal of politics* 77(3): 787–804.
- Kernell, Georgia. 2009. "Giving order to districts: Estimating voter distributions with national election returns." *Political Analysis* 17(3): 215–235.
- Key, Valdimer Orlando. 1955. "Politics, parties, and pressure groups."
- Key, V.O. Jr. 1949. "Southern politics in state and nation."
- Koger, Gregory, Seth Masket, and Hans Noel. 2009. "Partisan webs: Information exchange and party networks." *British Journal of Political Science*: 633–653.
- Kucukelbir, Alp et al. 2015. "Automatic variational inference in stan." In *Advances in neural information processing systems*, p. 568–576.
- La Raja, Raymond, and Brian Schaffner. 2015. *Campaign finance and political polarization: When purists prevail*. University of Michigan Press.

- Lax, Jeffrey R, and Justin H Phillips. 2009. "How should we estimate public opinion in the states?" *American Journal of Political Science* 53(1): 107–121.
- Layman, Geoffrey C., and Thomas M. Carsey. 2002. "Party Polarization and "Conflict Extension" in the American Electorate." *American Journal of Political Science* 46(4): 786. <http://www.jstor.org/stable/3088434?origin=crossref> (Accessed February 22, 2015).
- Layman, Geoffrey C et al. 2010. "Activists and conflict extension in american party politics." *American Political Science Review*: 324–346.
- Lebo, Matthew J, Adam J McGlynn, and Gregory Koger. 2007. "Strategic party government: Party influence in congress, 1789–2000." *American Journal of Political Science* 51(3): 464–481.
- Lelkes, Yphtach, and Paul M Sniderman. 2016. "The ideological asymmetry of the american party system." *British Journal of Political Science* 46(4): 825–844.
- Levendusky, Matthew. 2009. *The partisan sort: How liberals became democrats and conservatives became republicans*. University of Chicago Press.
- Levendusky, Matthew S, Jeremy C Pope, and Simon D Jackman. 2008. "Measuring district-level partisanship with implications for the analysis of us elections." *The Journal of Politics* 70(3): 736–753.
- Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe. 2009. "Generating random correlation matrices based on vines and extended onion method." *Journal of multivariate analysis* 100(9): 1989–2001.
- Liao, Shirley X, and Corwin M Zigler. 2020. "Uncertainty in the design stage of two-stage bayesian propensity score analysis." *Statistics in Medicine*.

- Link, William A., and Mitchell J. Eaton. 2011. "On thinning of chains in MCMC." *Methods in Ecology and Evolution* 3(1): 112–115. <https://doi.org/10.1111/j.2041-210x.2011.00131.x>.
- MacKay, David JC. 1992. "A practical bayesian framework for backpropagation networks." *Neural computation* 4(3): 448–472.
- Maisel, L Sandy, and Walter J Stone. 1997. "Determinants of candidate emergence in us house elections: An exploratory study." *Legislative Studies Quarterly*: 79–96.
- Mann, Thomas E. 1978. *Unsafe at any margin: Interpreting congressional elections*. Aei Pr.
- Martin, Andrew D, and Kevin M Quinn. 2002. "Dynamic ideal point estimation via markov chain monte carlo for the us supreme court, 1953–1999." *Political Analysis* 10(2): 134–153.
- Masket, Seth. 2009. *No middle ground: How informal party organizations control nominations and polarize legislatures*. University of Michigan Press.
- Mayhew, David R. 1974. *Congress: The electoral connection*. Yale University Press.
- Mayhew, David R. 1986. *Placing parties in american politics: Organization, electoral settings, and government activity in the twentieth century*. Princeton University Press.
- McCarty, Nolan, and Howard Poole Keith T. and Rosenthal. 2006. *Polarized america: The dance of ideology and unequal riches*. Cambridge, MA: MIT Press.
- McCarty, Nolan, Keith T Poole, and Howard Rosenthal. 2009. "Does gerrymandering cause polarization?" *American Journal of Political Science* 53(3): 666–680.
- McClosky, Herbert, Paul J Hoffmann, and Rosemary O'Hara. 1960. "Issue conflict and consensus among party leaders and followers." *The American Political Science Review* 54(2): 406–427.

- McGann, Anthony J. 2014. "Estimating the political center from aggregate data: An item response theory alternative to the stimson dyad ratios algorithm." *Political Analysis*: 115–129.
- McGhee, Eric et al. 2014. "A primary cause of partisanship? Nomination systems and legislator ideology." *American Journal of Political Science* 58(2): 337–351.
- Meager, Rachael. 2019. "Understanding the average impact of microcredit expansions: A bayesian hierarchical analysis of seven randomized experiments." *American Economic Journal: Applied Economics* 11(1): 57–91.
- Montgomery, Jacob M, Brendan Nyhan, and Michelle Torres. 2018. "How conditioning on posttreatment variables can ruin your experiment and what to do about it." *American Journal of Political Science* 62(3): 760–775.
- Neal, Radford M. 2012. "MCMC using hamiltonian dynamics." *arXiv preprint arXiv:1206.1901*.
- Norrande, Barbara. 1989. "Ideological representativeness of presidential primary voters." *American Journal of Political Science*: 570–587.
- Ornstein, Joseph T, and JBrandon Duck-Mayr. 2020. "Gaussian process regression discontinuity."
- Pacheco, Julianna. 2011. "Using national surveys to measure dynamic us state public opinion: A guideline for scholars and an application." *State Politics & Policy Quarterly*: 1532440011419287.
- Park, David K, Andrew Gelman, and Joseph Bafumi. 2004. "Bayesian multilevel estimation with poststratification: State-level estimates from national polls." *Political Analysis* 12(4): 375–385.
- Pearl, Judea. 1995. "Causal diagrams for empirical research." *Biometrika* 82(4): 669–688.

- Pearl, Judea. 2009. *Causality: Models, reasoning, and inference*. 2nd ed. Cambridge University Press.
- Petrocik, John Richard. 2009. "Measuring party support: Leaners are not independents." *Electoral Studies* 28(4): 562–572. <http://linkinghub.elsevier.com/retrieve/pii/S0261379409000511> (Accessed April 16, 2015).
- Phillips, Anne. 1995. *The politics of presence*. Clarendon Press.
- Pitkin, Hanna Fenichel. 1967. *The concept of representation*. Univ of California Press.
- Plummer, Martyn. 2015. "Cuts in bayesian graphical models." *Statistics and Computing* 25(1): 37–43.
- Poole, Keith T, and Howard Rosenthal. 1997. "Congress: A political-economic history of roll call voting." *New York: Oxford University Press*.
- Porter, Rachel A, and Sarah Treul. 2020. "Reevaluating experience in congressional primary elections."
- Rahn, Wendy M. 1993. "The role of partisan stereotypes in information processing about political candidates." *American Journal of Political Science*: 472–496.
- Ratkovic, Marc. 2019. "Rehabilitating the regression: Honest and valid causal inference through machine learning."
- Ratkovic, Marc, and Dustin Tingley. 2017. "Causal inference through the method of direct estimation." *arXiv preprint arXiv:1703.05849*.
- Robins, James M. 1997. "Causal inference from complex longitudinal data." In *Latent variable modeling and applications to causality*, Springer, p. 69–117.

- Rogowski, Jon C. 2016. "Voter decision-making with polarized choices." *British Journal of Political Science*: 1–22. <https://doi.org/10.1017%2Fs0007123415000630>.
- Rogowski, Jon C, and Stephanie Langella. 2015. "Primary systems and candidate ideology: Evidence from federal and state legislative elections." *American Politics Research* 43(5): 846–871.
- Rubin, Donald B. 1978. "Bayesian inference for causal effects: The role of randomization." *The Annals of statistics*: 34–58.
- Rubin, Donald B. 2005. "Causal inference using potential outcomes: Design, modeling, decisions." *Journal of the American Statistical Association* 100(469): 322–331.
- Samii, Cyrus, Laura Paler, and Sarah Zukerman Daly. 2016. "Retrospective causal inference with machine learning ensembles: An application to anti-recidivism policies in colombia." *Political Analysis* 24(4): 434–456.
- Shor, Boris, and Nolan McCarty. 2011a. "The ideological mapping of american legislatures." *American Political Science Review* 105(03): 530–551.
- Shor, Boris, and Nolan McCarty. 2011b. "The ideological mapping of american legislatures." *American Political Science Review* 105(03): 530–551.
- Sides, John et al. 2018. "On the representativeness of primary electorates." *British Journal of Political Science*: 1–9.
- Snyder, James M Jr. 1994. "Safe seats, marginal seats, and party platforms: The logic of platform differentiation." *Economics & Politics* 6(3): 201–213.
- Snyder Jr, James M. 1992. "Artificial extremism in interest group ratings." *Legislative Studies Quarterly*: 319–345.

- Stimson, James A. 1991. *Public opinion in america: Moods, cycles, and swings*. Westview Press.
- Stokes, Donald E. 1963. "Spatial models of party competition." *The American Political Science Review* 57(2): 368–377.
- Stone, Walter J, and L Sandy Maisel. 2003. "The not-so-simple calculus of winning: Potential us house candidates' nomination and general election prospects." *The Journal of Politics* 65(4): 951–977.
- Tausanovitch, Chris, and Christopher Warshaw. 2017. "Estimating candidates' political orientation in a polarized congress." *Political Analysis* 25(2): 167–187.
- Tausanovitch, Chris, and Christopher Warshaw. 2013. "Measuring constituent policy preferences in congress, state legislatures, and cities." *The Journal of Politics* 75(02): 330–342.
- Thomsen, Danielle M. 2014. "Ideological moderates won't run: How party fit matters for partisan polarization in congress." *The Journal of Politics* 76(3): 786–797.
- Thomsen, Danielle M. 2020. "Ideology and gender in us house elections." *Political Behavior* 42(2): 415–442.
- Tomz, Michael, and Robert P Van Houweling. 2008. "Candidate positioning and voter choice." *American Political Science Review*: 303–318.
- Treier, Shawn, and D Sunshine Hillygus. 2009. "The nature of political ideology in the contemporary electorate." *Public Opinion Quarterly* 73(4): 679–703.
- Treier, Shawn, and Simon Jackman. 2008. "Democracy as a latent variable." *American Journal of Political Science* 52(1): 201–217.
- Vansteelandt, Stijn. 2009. "Estimating direct effects in cohort and case–control studies." *Epidemiology*: 851–860.

- Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. “Practical bayesian model evaluation using leave-one-out cross-validation and waic.” *Statistics and computing* 27(5): 1413–1432.
- Vehtari, Aki et al. 2020. “Rank-normalization, folding, and localization: An improved \widehat{R} for assessing convergence of MCMC.” *Bayesian Analysis*.
- Wager, Stefan, and Susan Athey. 2018. “Estimation and inference of heterogeneous treatment effects using random forests.” *Journal of the American Statistical Association* 113(523): 1228–1242.
- Wang, Bo, and D Titterington. 2012. “Convergence and asymptotic normality of variational bayesian approximations for exponential family models with missing values.” *arXiv preprint arXiv:1207.4159*.
- Warshaw, Christopher, and Jonathan Rodden. 2012. “How should we measure district-level public opinion on individual issues?” *The Journal of Politics* 74(01): 203–219.