Do Primaries Work?

Nomination Politics and the Representation of Local Partisan Preferences

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, **IDK yet**

Barry C. Burden (Chair) ________________________________

Kenneth R. Mayer ________________________________

Eleanor Neff Powell ________________________________

Alexander M. Tahk ________________________________

5th member ________________________________

Note for APW

This document contains segments of work from my dissertation. The project is about primary nominations and the ideological representation of constituencies. I argue that research of U.S. elections is unclear about whether nomination contests serve their stated purpose: to represent the preferences of local partisan voters.

Political observers believe that nominating contests present candidates with a "strategic positioning dilemma"—to win both the nomination and the general election, candidates must position themselves as some ideological compromise between their partisan base and the district median voter. In Chapter 1, I argue that this view of primaries is theoretically and empirically tenuous. Theoretically, primary elections present voters with uniquely high informational demands that voters may not meet, especially when all candidates share the same party label. Moreover, informal party networks may overshadow voter preferences in shaping nomination outcomes. Empirically, research on primary representation is hindered by a limited understanding of local partisan preferences, which are crucial to the theory's view of intra-party representation. Many studies measure local preferences using rough proxies that do not distinguish between Republican and Democratic preferences, such as presidential vote shares, and many others do not operationalize citizen preferences at all.

In Chapter 2, I build an ideal point model for the policy ideology of the *median partisan* for Democratic and Republican constituents in every congressional district. All 435 congressional districts contain citizens who identify as Democrats and Republicans, and this model estimates the average preferences of Democratic and Republican identifiers as separate groups (870 ideal points, two per district). I sometimes refer to these 870 units as district-party publics or district-party groups. They are distinct from partisan primary *voters* or *caucus-goers*, who can be thought of as a subset of district-party publics (with some error). I refer to the measured construct as policy ideology, and I estimate it with survey data on policy preferences. It is distinct from concepts like "symbolic ideology."

I then apply these citizen ideology measures to test key claims inherent in the conventional view of primaries. Chapter 3 previews my approach for these empirical tests, which is a blend of causal inference and Bayesian estimation. If primaries function as we think they ought to, we should see that the level of conservatism (or liberalism) among district-party groups influences the conservatism (liberalism) of candidates who run for party nominations in those districts. This is the focus of Chapter 4. We should also see that the policy ideology of the district-party group affects which candidate from a menu of options is nominated, which is the focus of Chapter 5.

The material in this current document is a selection from Chapters 2 and 4, with some pages skipped. This breaks some of the smart references to subsections, so I apologize for random ?? characters that appear. If you would like background info on understanding IRT models, you can read straight through. If you are familiar with IRT models and want to skip to learning about *my* model, you can skip to Section 2.2. If you would rather not read about the guts of my model and instead skip to the empirical application, that begins in Chapter 4! I appreciate your patience with this project-in-progress.

Contents

List of Tables

List of Figures

— 1 —

Primary Elections and Ideological Choice

Dropped for APW.

— 2 —

Modeling the Partisan Constituency's Policy Preferences

To study how partisan constituencies are represented in primary elections, we require a measure of the partisan constituency's policy preferences. This chapter presents the statistical model that I use to estimate the policy ideal points of district-party publics.

This chapter proceeds in three major steps. First, I review the theoretical basis for ideal point models, which can be traced to spatial models of policy choice from classic formal theory work in American political science such as Downs (1957). I connect these formal models to statistical models of policy ideal points (in a style that follows Clinton, Jackman, and Rivers 2004) as well as their connection to Item Response Theory (IRT) models from psychometrics and education testing (e.g. Fox 2010).

Second, I specify and test the group-level model that I build and employ in my analysis of district-party publics. This discussion includes details that are relevant to Bayesian estimation, including identification restrictions on the latent policy space, specification of prior distributions, and model parameterizations that expedite estimation with Markov chain Monte Carlo (MCMC). I begin with a static model for one time period, and then I describe a dynamic model that smooths estimates across time using hierarchical priors for model parameters (Caughey and Warshaw 2015). I test both the static and the dynamic models by fitting them to simulated data and determining how well they recover known parameter values.

Lastly, I describe how I fit the model to real data. This section describes data collection, data processing, and model performance, and a descriptive analysis of the estimates.

2.1 Material omitted here

2.1.1 The "Item Response Theory" Approach to Survey Response

Scholars of ideal point models have noted their similarity to models developed under item response theory (IRT) in psychometrics (for example, Londregan 1999). IRT models have a similar mission as ideal point models: measuring latent features in the data given individuals' response patterns to various stimuli. The canonical psychometric example is in education testing, where a series of test questions is used to measure a student's latent academic "ability" level. This section connects ideal point models to IRT in order to explain their important theoretical and mathematical intuitions.

2.1.1.1 Latent Traits

The first important feature to note about IRT models is that they are *measurement models*. The goal of a measurement model is to use observed data $\mathbf{y}$ to estimate some construct of theoretical interest $\boldsymbol{\theta}$, supposing that there is a distinction between the two. The observed data $\mathbf{y}$ are affected by $\boldsymbol{\theta}$, but there is no guarantee of a one-to-one correspondence between the two because $\boldsymbol{\theta}$ is not directly observed. We can represent a measurement model with general notation $\mathbf{y} = f(\boldsymbol{\theta}, \boldsymbol{\sigma})$, where $\boldsymbol{\sigma}$ represents some vector of auxiliary model parameters to be estimated in addition to $\boldsymbol{\theta}$ by fitting the model to observed data.

In an educational testing context, students take standardized tests intended to measure their academic "ability" levels. Analysts who score the tests cannot observe a student's ability directly—it is unclear how that would be possible. They do, however, observe the student's answers to known test questions. IRT models provides a structure to infer abilities from the student's pattern of test answers. The context of policy choice is similar. It is impossible to observe any individual's political ideology directly, but we theorize that it affects their responses to survey items about policy choices.

The IRT setup lets us summarize an individual's policy preferences by analyzing the structure of their responses to various policy choices.

It is crucial to note that the only way to estimate a latent construct from observed data is for the model to impose assumptions about the functional relationship between the latent construct and the observed data. In this sense, the estimates can be sensitive to the model's assumptions. While this is always important to acknowledge, it is also valuable to note that model-dependence is an ever-present consideration even for simpler measurement strategies, such as an additive index that sums or averages across a battery of variables. In fact, additive indices are special cases of measurement model where key parameters are assumed to be known and fixed, which is problematic if there is any reason to suspect that item responses are correlated across individuals.[1] In this way, measurement models *relax* the assumptions of simpler measurement strategies, even if the underlying mathematics are more intensive.

2.1.1.2 Item Characteristics and Item Parameters

Measurement models relax assumptions about the data's functional dependence on the construct of interest. Item response theory focuses this effort on the items to which subjects respond. Different items may reveal different information about the latent construct; the design of the model governs how those item differences can manifest (see Fox 2010 for a comprehensive review of IRT modeling).

Consider a simple model where a student i is more likely to answer test questions j correctly if she has greater academic ability θ_i. Analogously, a citizen who is more conservative is more likely to express conservative preferences for policy question j. Keeping the probit functional form from above, we can represent this simple model with the equation:

$$\Pr\left(y_{ij} = 1\right) = \Phi\left(\theta_i\right), \tag{2.1}$$

where θ_i is scaled such that the probability of a correct/conservative response is 0.5 at $\theta_i = 0$. This model makes the implicit assumption that knowing θ_i is sufficient to produce exchangeable response

[1] Midpoint and discrimination parameters would be sources of this correlation. Additive indices are similar to a model where all all midpoint and discrimination are respectively equal to 0 and 1 by assumption.

data; there are no systematic differences in the difficulty level of the test questions or the ideological nature of the policy choices that would affect the propensity of subjects to answer correctly/conservatively on average. This implicit assumption is often unrealistic. Just as some test questions are naturally more difficult than others, some policy questions present more extreme or lopsided choices than others, leading citizens with otherwise equivalent θ values to vary systematically in their response probability across items. Although the "ability-only" model seems unrealistic when posed as such, political science is replete with additive measurement scales that omit all item-level variation: indices of policy views, the racial resentment scale, survey-based scales of political participation, and more.

Rather than assume that all items behave identically for all individuals, IRT explicitly models the systematic variation at the item level using *item parameters.* IRT models have different behaviors based on the parameterization of the item effects in the model. The simplest IRT model is the "one-parameter" model,[2] which includes an item-specific intercept κ_j.

$$\Pr\left(y_{ij} = 1\right) = \Phi\left(\theta_i - \kappa_j\right) \tag{2.2}$$

IRT parlance refers to the κ_j parameter as the item "difficulty" parameter. In the testing context, if a student has higher ability than the difficulty of the question, the probability that they answer the test item correctly is greater than 0.5. This probability goes up for students with greater ability relative to item difficulty, and it goes down for items with greater difficulty relative to student ability. In a policy choice context, the difficulty parameter is better understood as the "cutpoint" parameter, the midpoint between two policy choices where the respondent is indifferent between the choice of Left or Right on item j. These cutpoints are allowed to vary from item to item; some policy choices present alternatives that are, on average, more conservative or liberal than others. For instance, the choice of *how much* to cut capital gains taxes will have a more conservative cutpoint than a question of whether to cut capital gains taxes at all. If there were no systematic differences across items, it would be the case that $\kappa_j = 0$ for all j, and the one-parameter model would reduce to the simpler model in Equation (2.1).

The "two-parameter" IRT model is more common, especially in the ideal point context. The two-parameter model introduces the "discrimination" parameter ι_j, which behaves as a slope on the

[2] One-parameter logit models are often called "Rasch" models, whereas their corresponding probit models are often called "Normal Ogive" models (Fox 2010).

difference between θ_i and κ_j.

$$\Pr\left(y_{ij} = 1\right) = \Phi\left(\iota_j\left(\theta_i - \kappa_j\right)\right), \tag{2.3}$$

Intuitively, the discrimination parameter captures how well a test item differentiates between the responses of high- and low-ability students, with greater values meaning more divergence in responses. In the ideal point context, it captures how strongly a policy question divides liberal and conservative respondents.[3]

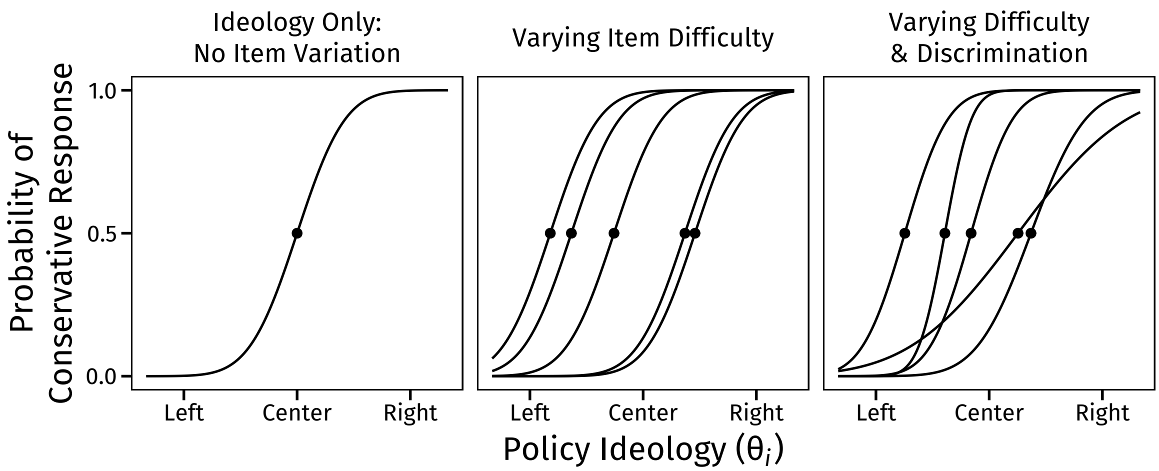

Figure 2.1: Examples of item characteristic curves under different item parameter assumptions

Figure 2.1 shows how response probabilities are affected by the parameterization of item effects. Each panel plots how increases in subject ability or conservatism (the horizontal axis) result in increased response probability (the vertical axis), where the shape of the curve is set by values of the item parameters. These curves are commonly referred to as *item characteristic curves* (ICCs) or *item response functions* (IRFs). The leftmost panel shows a model with no item effects whatsoever; any item is theorized to behave identically to any other item, and response probabilities are affected only by the subject's ability (ideology). The middle panel shows a one-parameter model where item difficulties (cutpoints) are allowed to vary systematically at the item level. Difficulty parameters behave as intercept shifts, so they convey which value of θ yields a correct response with probability 0.5, but

[3]Two-parameter IRT models are sometimes written with ι_j is distributed through the equation: $\iota_i\theta_i + \alpha_j$, where $\alpha_j = \iota_j\kappa_j$. Although this parameterization more closely follows a linear slope-intercept equation, it loses the appealing interpretation of κ_j as the midpoint between policy choices.

they do not affect the *elasticity* of the item response function to changes in θ. The final panel shows item response functions from the two-parameter IRT model, where item difficulties (intercepts) and discriminations (slopes) are allowed to vary across items.

2.1.1.3 IRT Interpretation of the Ideal Point Model

How do we interpret our statistical model of ideal points in light of item response theory? Recall the statistical model that we derived from the utility model above. An Actor i faces policy question j, with a Right alternative located at R_j and a Left alternative located at L_j. The Actor chooses the alternative closest to her ideal point θ_i, subject to idiosyncratic utility shocks summarized by ε_{ij}. Letting y_{ij} indicate the outcome that Actor i chooses the Right position on policy question j, the probability that $y_{ij} = 1$ is given by the two-parameter model in Equation (2.3) (or (**??**) above).

The behavior of the item parameters can be understood by remembering that they are functions of the Left and Right choice locations. For instance, the cutpoint parameter κ_j represents an intercept shift for an items response function and is equal to $\frac{R_j + L_j}{2}$. Suppose that $\theta_i - \kappa_j = 0$, which occurs if the item cutpoint falls directly on an Actor's ideal point. In such a case, the Actor would be indifferent (in expectation) to the choice of Left or Right, and the probability of choosing Right would collapse to 0.5.[4] The value of κ_j increases by moving either the Right or Left alternatives to the right (increasing R_j or L_j), subject to the constraint that $R_j \geq L_j$. Larger values of the item cutpoint imply a lower probability that the Actor chooses Right, since κ_j has a non-positive effect on the conservative response probability.[5] The opposite intuition holds as the Left position becomes increasingly progressive, resulting in larger values of κ_j that imply a higher probability of choosing Right, all else equal.

The discrimination parameter behaves as a "coefficient" on the distance between the Actor ideal point and the cutpoint, meaning that the Actor's choice is more elastic to her policy preferences as ι_j increases.[6] Because $\iota_j = 2(R_j - L_j)$, the discrimination parameter grows when the distance between

[4] This holds in logit and probit models, since $\text{logit}^{-1}(0)$ and $\Phi(0)$ are both equal to 0.5.

[5] Formally we can show this by taking the derivative of the link function with respect to the cutpoint: $\frac{\partial \iota_j (\theta_i - \kappa_j)}{\partial \kappa_j} = -\iota_j$, where ι_j is constrained to be greater than or equal to zero

[6] Again we can demonstrate this by noticing that the derivative of the link function with respect to the discrimination parameter is $\frac{\partial \iota_j (\theta_i - \kappa_j)}{\partial \iota_j} = (\theta_i - \kappa_j)$. The derivative's magnitude depends on the absolute value of this distance, and its

the Right and Left alternatives grows larger, which happens when R_j increases or L_j decreases.

In a special case that Right and Left alternatives are located in exactly the same location, the result is $\kappa_j = \iota_j = 0$, leading all Actors to choose Right with probability 0.5. This result represents a situation where policy preferences are not systematically related to the choice whatsoever, and only idiosyncratic error affects the choice of Right or Left. Although the model implies that this result is *mathematically* possible, it is not realistic to expect any of the policy choices in this project to induce this behavior.

2.1.2 Applications of the IRT Approach

A section reviewing IRT in political science:

- ideal points (Poole and Rosenthal, CJR, Londregan, Jeff Lewis, Michael Bailey, Martin and Quinn)
- citizen ideology with survey data (Seth Hill multinomial and individual, Tausanovitch and Warshaw small area, Caughey and Warshaw groups within areas)
- other latent modeling (Levendusky, Pope, and Jackman 2008)

2.2 Modeling Party-Public Ideology in Congressional Districts

This section outlines my group-level ideal point model for party publics. It begins by describing the connection between the individual-level IRT model and the group-level model and its implication for the parameterization of the model (Section 2.2.1). I then lay out the hierarchical model for party-public ideal points in its static form (Section 2.2.2) and its dynamic form (Section **??**). I discuss technical features of model implementation, including choices for model parameterization, model identification, prior distributions, and model testing methods such as prior predictive checks and posterior predictive checks.

sign depends on the sign of the difference.

2.2.1 Group-Level IRT Model

So far we have modeled individual responses to policy items according to their own individual ideal points, but this project is concerned with the average ideal point of a *group* of individuals. In the group model, we assume that individual ideal points are distributed within a group g, where groups are define as the intersection of congressional districts d and political party affiliations p.

As before, we observe a binary response from individual i to item j, which we regard as a probabilistic conservative response with probability π_{ij}, which is given a probit model.

$$y_{ij} \sim \text{Bernoulli}(\pi_{ij}) \tag{2.4}$$

$$\pi_{ij} = \Phi\left(\iota_j\left(\theta_i - \kappa_j\right)\right) \tag{2.5}$$

Following Fox (2010) and Caughey and Warshaw (2015), it is helpful to reparameterize the IRT model to accommodate a group-level extension. This parameterization replaces item "discrimination" with item "dispersion" using the parameter $\sigma_j = \iota_j^{-1}$ and rewriting the model as

$$\pi_{ij} = \Phi\left(\frac{\theta_i - \kappa_j}{\sigma_j}\right). \tag{2.6}$$

Caughey and Warshaw (2015) describe the dispersion parameter as introducing "measurement error" in π_{ij} beyond the standard Normal utility error from ε_{ij} above.

The group model begins with the notion that there is a probability distribution of ideal points within a group g, where a "group" is a partisan constituency within a congressional district. Supposing that individual deviations from the group mean are realized by the accumulation of a large number of random forces, we can represent an individual ideal point as a Normal draw from the group,[7]

$$\theta_i \sim \text{Normal}\left(\bar{\theta}_{g[i]}, \sigma_{g[i]}\right) \tag{2.7}$$

where $\bar{\theta}_{g[i]}$ and $\sigma_{g[i]}$ are the mean and standard deviation of ideal points within i's group g.

While it is possible to continue building the model hierarchically from (2.7), it would be far too computationally expensive to estimate every individual's ideal point in additional to the group-level parameters—every individual ideal point is essentially a nuisance parameter. Instead, we rewrite the

[7] Notation for Normal distributions will always describe the scale parameter in terms of standard deviation σ instead of variance σ^2. This keeps the notation consistent with the way Normal distributions are expressed in Stan code.

model by aggregating individual-level survey response data to the group level, expressing the grouped outcome data as a function of the group parameters. Let $s_{gj} = \sum_{i \in g}^{n_{gj}} y_{ij}$, the number of conservative responses from group g to item j, where n_{gj} is the total number of responses (trials) to item j by members of group g. Supposing these trials were collected independently across groups and items (an assumption that is relaxed later), we could model the grouped outcome as a binomial random variable,

$$
\begin{aligned}
s_{gj} &\sim \text{Binomial}\left(n_{gj}, \bar{\pi}_{gj}\right) \\
\bar{\pi}_{gj} &= \Phi\left(\frac{\bar{\theta}_g - \kappa_j}{\sqrt{\sigma_g^2 + \sigma_j^2}}\right),
\end{aligned}
\tag{2.8}
$$

where $\bar{\pi}_{gj}$ is the "average" conservative response probability for item j in group g, or the probability that a randomly selected individual from group g gives a conservative response to item j. Our uncertainty about any random individual's ideal point relative to the group mean is included in the model as group-level variance term. If individual ideal points are Normal within their group, this within-group variance can simply be added to the probit model as another source of measurement error, with larger within-group variances further attenuating $\bar{\pi}_{ij}$ toward 0.5. Caughey and Warshaw (2015) derive this result in the supplementary appendix to their article.

The current setup assumes that every item response is independent, conditional on the group and the item. This assumption is violated if the same individuals in a group answer multiple items—one individual who answers 20 items is less informative about the group average than 20 individuals who answer one item apiece. While this too could be addressed by explicitly modeling each individual's ideal point (extending the model directly from Equation (2.7)), I implement a weighting routine that downweights information from repeated-subject observations while adjusting for nonrepresentative sample design, as I will describe in Section **??**.

2.2.2 Hierarchical Model for Group Parameters

The group model described so far can be estimated straightforwardly if there are enough responses from enough individuals in enough district-party groups. In practice, however, a single survey will not contain a representative sample of all congressional districts, and certainty not a representative sample of partisans-within-districts. I specify a hierarchical model for the group parameters in order to stabilize the estimates in a principled way. The hierarchical model learns how group ideal points are related to cross-sectional (and eventually, over-time) variation in key covariates, borrowing strength from data-rich groups to stabilize estimates for data-sparse groups, and even imputing estimates for groups with no survey data at all. This section describes the multilevel structure using traditional notation for hierarchical models; later in Section **??** I describe how I parameterize the model for the estimation routine.

I posit a hierarchical structure where groups g are "cross-classified" within districts d and parties p. This means that groups are nested within districts and within parties, but districts and parties have no nesting relationship to one another. Districts are further nested within states s. I represent this notationally by referring to group g's district as $d[g]$, or the g^{th} value of the vector $\mathbf{d}$. Similarly, g's party is $p[g]$. For higher levels such as g's state, I write $s[g]$ as shorthand for the more-specific but more-tedious $s[d[g]]$.

I use this hierarchical structure to model the probability distribution of group ideal points $\bar{\theta}_g$. I consider the group ideal point as a Normal draw from the distribution of groups whose hypermean is predicted by a regression on geographic-level data with parameters that are indexed by political party. This regression takes the form

$$\bar{\theta}_g \sim \text{Normal}\left(\mu_{p[g]} + \mathbf{x}_{d[g]}^{\top}\beta_{p[g]} + \alpha_{s[g]p[g]}^{\texttt{state}}, \sigma_p^{\texttt{group}}\right) \tag{2.9}$$

where $\mu_{p[g]}$ is a constant specific to party p,[8] $\mathbf{x}_d$ is a vector of congressional district-level covariates with party-specific coefficients β_p. State effects $\alpha_{sp}^{\texttt{state}}$ are also specific to each party. The benefit of specifying separate parameters for each party is that geographic features (such as racial composition, income inequality, and so on) may be related to ideology in ways that are not identical across all parties. This is an important departure from the structure laid out by Caughey and Warshaw (2015), which estimates the same set of geographic effects for all groups in the data.

The state effects are regressions on state features as well,

$$\alpha_{sp}^{\texttt{state}} \sim \text{Normal}\left(\mathbf{z}_s^{\top}\gamma_p + \alpha_{r[s]p}^{\texttt{region}}, \sigma_p^{\texttt{state}}\right), \tag{2.10}$$

[8] Or "grand mean," since all covariates are eventually be centered at their means.

where state-level covariates $\mathbf{z}_s$ have party-specific coefficients γ_p. Each state effect is a function of a party-specific region effect $\alpha^{\texttt{region}}_{[s]rp}$ for Census regions indexed r, which is a modeled mean-zero effect to capture correlation within regions.

$$\alpha^{\texttt{region}}_{rp} \sim \text{Normal}\left(0, \sigma^{\texttt{region}}_{p}\right) \tag{2.11}$$

2.2.3 Identification Restrictions

Ideal point models, as with all latent space models, are unidentified without restrictions on the policy space. The model as written can rationalize many possible estimates for the unknown parameters, with no prior basis for deciding which estimates are best. A two-parameter model such as this requires some restriction on the polarity, location, and scale of the policy space.

- Location: the latent scale can be arbitrarily shifted right or left. We could add some constant to every ideal point, and the response probability would be unaffected if we also add the same constant to every item cutpoint.
- Scale: the latent scale can be arbitrarily stretched or compressed. We could multiply the latent space by some scale factor, and the response probability would be unaffected if we also multiply the discrimination parameter by the inverse scale factor.
- Polarity: the latent scale could be reversed. We could flip the sign of every ideal point, and the response probability would be unaffected if we also flip the sign of every item parameter.

These properties are present with every statistical model, but covariate data typically provide the restrictions necessary to identify a model.[9] Because the response probability is a function of the interaction of multiple parameters in a latent space, however, data alone do not provide the necessary restrictions on the space to provide a unique solution. Absent any natural restriction from the data, I provide my own restrictions on the polarity, location, and scale of the policy space.

The polarity of the space is fixed by coding all items such that conservative responses are 1 and liberal responses are 0. This ensures that increasing values on the link scale always lead to an increasing

[9] We could imagine shifting, stretching, or reversing the sign of a covariate to reveal the same mathematical behaviors. All of these transformations would result in the same predictions as long as the parameters are also transformed to compensate.

probability of a *conservative* item response. Additionally I impose a restriction that all discrimination parameters are positive, which implies that shifting any ideal point farther to the *right* of an item cutpoint increases the probability of a conservative response, all else equal.

The location of the space is set by restricting the sum of the J item cutpoints to be 0. If $\tilde{\kappa}_j$ were an unrestricted item cutpoint, the restricted cutpoint κ_j used in response model would be defined as

$$\kappa_j = \tilde{\kappa}_j - \frac{\sum\limits_{j=1}^{J} \tilde{\kappa}_j}{J}, \tag{2.12}$$

which is performed in every iteration of the sampler. This restriction on the sum of the cutpoint parameters also implies a restriction on the mean of the cutpoints, since $\frac{0}{J} = 0$.

Lastly, I set the scale of the latent space by restricting the product of the J discrimination parameters to be equal to 1. I implement this by restricting the log discrimination parameter to have a sum of 0, which achieves an equivalent transformation.[10] Letting $\tilde{\iota}_j$ be the unrestricted discrimination parameter, we obtain the restricted ι_j as follows.

$$\iota_j = \exp\left(\log\left(\iota_j\right)\right) \tag{2.13}$$

$$\log\left(\iota_j\right) = \log\left(\tilde{\iota}_j\right) - \frac{1}{J}\sum_{j=1}^{J} \log\left(\tilde{\iota}_j\right) \tag{2.14}$$

Item discrimination is then reparameterized as dispersion, $\sigma_j = \iota_j^{-1}$. These restrictions on the item parameters are sufficient to identify $\bar{\theta}_g$.

2.2.4 Material omitted

[10] A quick demonstration using three unknown values a, b, and c. If $a \times b \times c = 1$, then $\log(a) + \log(b) + \log(c) = \log(1)$, which is equal to 0.

— 3 —

Bayesian Approach to Causal Inference in Political Science

Dropped for APW.

— 4 —

Candidate Positioning in Primary Elections

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 candidates who run to represent that district-party.

It is important to distinguish the influence of the district-party public from the influence of district *partisanship*. Does Senator Susan Collins (R-ME) have a reputation as a moderate Republican senator because the balance of a close numerical balance of Republican and Democratic voters in Maine? Or is it because her Republican constituency in Maine is relatively moderate compared to the Republican constituencies in other states represented by more conservative Senators? 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 district construct from aggregate voting patterns. 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, because district-level voting behavior is certainly affected by district-party ideology, and it likely mediates the effect of district-party ideology on candidate positioning. Estimating the direct effect of district-party ideology by simply controlling for district voting in a regression is likely to introduce collider bias by conditioning on a post-treatment variable (Greenland, Pearl, and Robins 1999).

This chapter investigates the effect of district-party ideology on primary candidate positions using a sequential-g model, a multistage approach that estimates the direct effect of district-party ideology while holding fixed district-level voting, a likely mediator of the total effect (Acharya, Blackwell, and Sen 2016). I use causal graphs to illustrate the modeling problem and discuss the assumptions required for identifying the targeted effect, which I adapt to a multilevel context where treatment and mediators exist at different levels of the data hierarchy. Finally, I describe and implement a method for propagating measurement uncertainty through the sequential-g method, since the key independent variable is an uncertain estimate from a measurement model.

4.1 Constituency Preferences and Candidate Positioning

Lit review:

- I have lit review but it isn't well organized right now. I'll summarize some main points:
- classic models of electoral competition view candidate moderation as an electoral benefit. You can make non-moderate positions make sense if you incorporate campaign volunteers, primaries, etc.
- Researchers find evidence that district-level aggregate voting (e.g. the presidential vote) is related to candidate positioning.
- Why do candidates take non-moderate stances in reality? Studies that look into the polarizing effects of primaries find little evidence that primaries matter for candidate positioning, but many of these studies are focused on Congressional incumbents.
- Studies that include non-incumbent candidates show that many of the empirical implications of the "strategic positioning dilemma" don't receive a lot of evidence. For instance, studies of primary "openness" consistently show basically zero or even reversed relationships to candidate ideology as you would expect from theory.

With direct measures of district-party policy ideology, do we get a different topline picture of within-party representation in primaries? Figure 4.1 shows a descriptive picture of the relationship between district-party ideology and candidate ideology for House primary candidates, the latter measured with dynamic ideal point scores derived from campaign contributions (Bonica 2014). Across incumbents, challengers, and open-seat candidates, in both parties, and across several election cycles, we observe a generally positive relationship between these two measures. As partisan citizens in a district become more conservative, so too do the candidates who run in those districts.

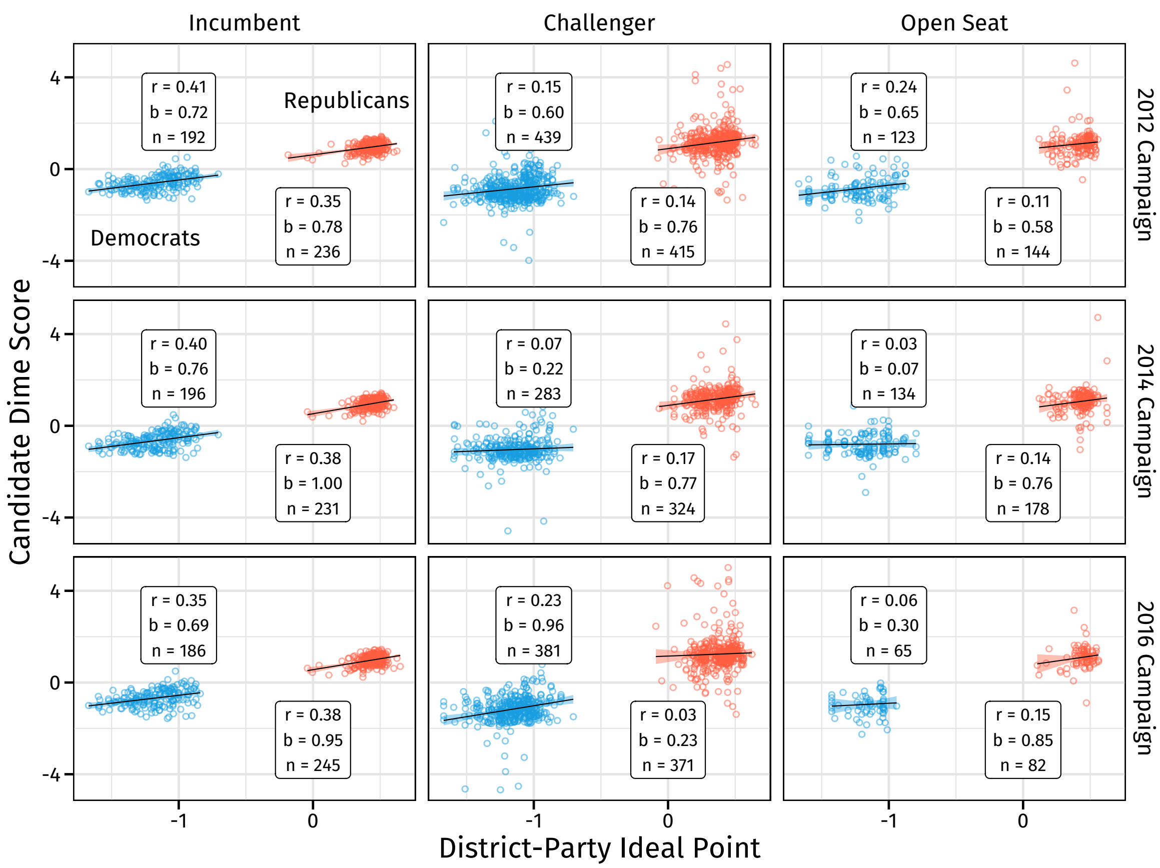

Figure 4.1: Simple comparison of candidate Dime scores against their respective district-party ideal points. Results generally indicate a positive relationship: as partisan constituents in a district hold more conservative policy preferences, candidates running for that party's nomination in the House are also more conservative.

The strength of these relationships vary most dramatically by candidate incumbency status, which could be owed to competitive positioning incentives when, for example, challengers attempt to differentiate themselves ideologically from incumbents in ways that aren't entirely explained by bottom-up ideological pressure from voters (c.f. Ansolabehere, Snyder, and Stewart 2001; Burden 2004). Incumbent candidates' ideal points are most strongly related to district-party ideology, consistent with a notion that ideological consistency creates a selection effect into incumbency in previous election.

This descriptive picture informs a few modeling choices during the sequential-g routine. Because incumbency status appears to be a substantial modifier of the relationship between citizen and candidate ideology, I estimate the direct effect of district-party ideology separately for incumbents, challengers, and open seat candidates. By comparison, estimates are quite similar across election cycles, so I pool election cycles into one model with cycle fixed effects rather than estimating entirely separate models for different cycles. And although the descriptive relationships vary modestly across parties, the variables that could confound the relationship between citizen and candidate ideology could differ dramatically across parties. I therefore estimate models for Democrats and Republicans separately. This results in six groups of sequential-g estimates: three incumbency categories × two major parties.

Although this descriptive picture is suggestive about the relationship between district-party ideology and candidate ideology, we need more rigorous methods to interrogate a causal relationship. In particular, we should be concerned that district features that promote conservative voters also promote conservative candidates, so background features of districts are important to control. Furthermore, if it is worthwhile for researchers to consider the unique effect of district-party ideology, it is important to demonstrate that it affects candidate positioning above and beyond its intermediate effect on district voting. The sequential-g approach confronts both of these threats to inference.

4.2 The Direct Effect of Partisan Preferences

This section describes the causal estimand and estimation routine that follows. Sequential-g estimates a quantity called the *average controlled direct effect*, the average effect of a treatment on an outcome, holding fixed a mediator variable for all units under consideration.

Consider a potential outcome where candidate positioning C is affected by district voting V and district-party ideology T, or $C(T,V)$. The controlled direct effect imagines that we can intervene on both T and V, varying the value of T between t and t' while fixing $V = v$. The controlled direct effect is defined for a single unit i as,

$$CDE_i(t,t',v) = C_i(t,v) - C_i(t',v), \tag{4.1}$$

or, how would C_i change if we could vary t without influencing v in the process? The dependence of district voting V on district-party ideology T is shown by the causal graph in Figure 4.2. A model that estimates the *total* effect of district-party ideology will fail to differentiate the fraction of the effect flowing through path $T \rightarrow C$ from the fraction of the effect through path $T \rightarrow V \rightarrow C$. However, simply controlling for V will not isolate the direct effect, since it can open back-door paths from T to C through confounders represented by U (Montgomery, Nyhan, and Torres 2018). If there are variables that affect aggregate district voting that are independent of district-party ideology, such as valence features from unrelated prior candidates, post-treatment conditioning on the district vote can create confounders unintentionally. Sequential-g is a special case of a broader class of models (structural nested mean models) that measure direct effects by subtracting intermediary effects without creating collider bias (Acharya, Blackwell, and Sen 2016; Vansteelandt 2009).

In order to implement a sequential-g routine, we need to specify valid models that separately identify the mediator-outcome relationship (the total effect of district voting on candidate positioning) and the treatment-outcome relationship (the total effect of district-party ideology on candidate positioning). This twin identification is formalized using an assumption of *sequential ignorability*, or sequential unconfoundedness (Robins and Greenland 1994). This means that unit potential outcomes $C_i(t,v)$ are independent of treatment, conditional on pre-treatment covariates X_i,

$$C_i(t,v) \perp\!\!\!\perp T \mid X_i = x \tag{4.2}$$

and secondly that potential outcomes are independent of the mediator, conditional on treatment, pre-treatment covariates, and *intermediate covariates* Z_i that may affect the mediator separately from T and X.

$$C_i(t,v) \perp\!\!\!\perp M \mid T_i = t, X_i = x, Z_i = z \tag{4.3}$$

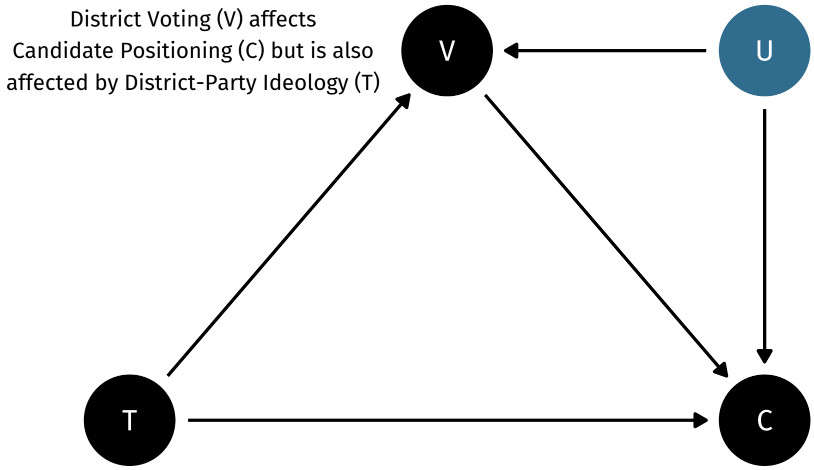

Figure 4.2: A DAG that presents district partisanship as a collider along the path from district-party preferences to candidate positioning. Controlling for district partisanship can bias the causal estimate of district-party preferences in several ways. If there is a path $\theta \rightarrow P \rightarrow Y$, then the effect of θ does not represent the total effect. In the presence of intermediate confounder U, controlling for district partisanship induces collider bias by unblocking the path $\theta \leftarrow U \rightarrow Y$.

Figure 4.3 visualizes the modeling assumptions for sequential-g estimation using causal graphs, which helps explain how to implement the routine. The left panel shows the stage-one model, which estimates the effect of past voting V (the mediator) on candidate positioning C (the outcome). This first stage conditions on district-party preferences T, pre-treatment confounders X, and intermediate confounders Z, all of which are necessary to identify the causal effect of the mediator.

After estimating the mediator's effect on the outcome, the outcome variable is *demediated* by subtracting the mediator effect from the outcome variable. The center panel represents this demediation step by rewriting the outcome variable as $b(C)$, the demediated value of C.[1] The stage-two model then estimates the effect of district-party preferences T on the demediated candidate positions, controlling for pre-treatment covariates X. Demediating the outcome suppresses the path from V to $b(C)$ because (by sequential unconfoundedness) there is no longer any systematic variation between the mediator and the outcome. As such, there is no need to adjust for V, since it has no systematic effect on $b(C)$ after demediation. Furthermore, although there remains a causal effect from the intermediate confounders Z to candidate positions, the stage-two model does not adjust for these

[1]The exact demediation operation is shown below in Equation (4.5).

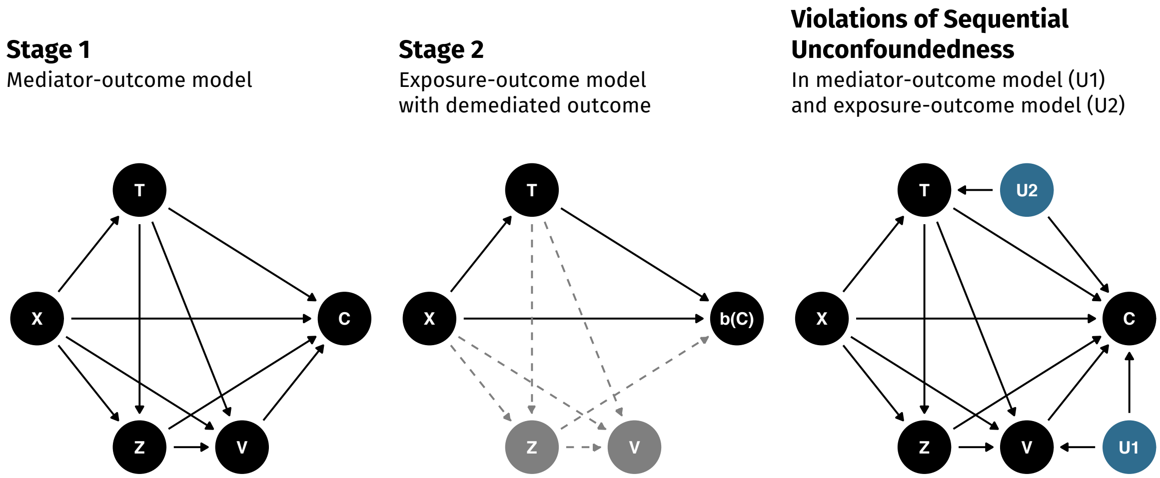

Figure 4.3: Directed acyclic graphs that describe the modeling problem and sequential-g estimation. The stage 1 mediator model estimates the effect of the past presidential vote (V) on candidate positioning (C). The stage 2 exposure model subtracts the vote effect from candidate positions and estimates the effect of partisan group ideology (T) on the demediated candidate positions ($b(C)$). The final panel shows where unadjusted confounders violate the identification assumptions in stage 1 ($U1$) and in stage 2 ($U2$).

confounders to avoid post-treatment bias in the estimate of the CDE. This stage-two model recovers the controlled direct effect of T on C.

It is worth noting here that if we estimate the stage-two model using the natural value of C rather than the demediated $b(C)$, we would obtain the total effect of T (the conditional average treatment effect) rather than the controlled direct effect. This can be valuable, since the difference between the total effect and the controlled direct effect are an indirect indicator of how much of the total effect flows through a mediator variable.

The final panel shows where unmeasured confounding can violate the sequential unconfoundedness assumption. The stage-one model identifies the causal effect of the mediator, so if an unmeasured variable (represented in the future by $U1$) affects both the mediator and the outcome, the mediator's effect is not identified. Similarly, the stage-two model does not identify the effect of T on $b(C)$ if they share an unmeasured confounder $U2$. Unmeasured variables in other locations of the graph certainly exist, but they do not violate sequential unconfoundedness unless they can be represented by an open back-door path through $U1$ or $U2$.[2] There are functional form considerations however—biases may

[2]For instance, a variable W may be a common cause of X and C, thus creating the path $T \to X \to W \to C$, but it does

remain if the functional forms for confounders X and Z are inappropriate.[3]

4.2.1 Multilevel Considerations

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.

That setup violates likely violates the stable unit treatment value assumption (SUTVA) within a given group—even if candidates in the same group receive the same treatment, the positioning of one candidate may affect the positioning of another, violating 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 averaging across groups (Hill n.d.). In potential outcomes notation, even if we can define potential outcomes at the individual level ($Dime_i(\bar{\theta}_{g[i]})$), the lowest level where we could credibly identify treatment effects would be at the group level, where the potential outcome for a group is the average outcome within the group ($\overline{Dime}_g\left(\bar{\theta}_g\right)$).

4.2.2 Bayesian Considerations

One important feature of this dissertation is that our causal variable of interest, the policy ideology of district-party publics, is not measured with certainty. The model in Chapter 2 estimates this quantity up to a posterior distribution, but the variation in the posterior distribution creates an additional layer of uncertainty in our understanding of causal effects. This creates theoretical issues with the definition of treatment effects as well as practical issues for estimation, which I describe in turn.

As discussed in Chapter 3, it is natural to interpret causal inference as Bayesian updating about a causal parameter using models and data. Following Rubin (1978)'s Bayesian interpretation of his own

not confound the effect of T because conditioning on X blocks the path. Additionally, although intermediate confounders Z are presented as descendants of group policy ideology T, intermediate confounders do not necessarily have to be affected by T (though they can be). They only need to be confounders of the mediator-outcome relationship.

[3] I plan to investigate more flexible functional forms in the near future.

work with potential outcomes, causal inference is a method for generating counterfactual estimates of an outcome variable as if they were predictive draws from the posterior distribution. Supposing that we seek the causal effect of an intervention on group ideology, τ, and we observe only one potential outcome, $y_g(\bar{\theta}_g)$, how do we specify the posterior distribution if we set group ideology to some other value θ'? Not only do we have posterior uncertainty about τ, but we are also uncertain about the actual value of $\bar{\theta}_g$ that we observed in the first place, since the measurement model from Chapter 2 estimates it only up to a posterior distribution. Our posterior distribution for missing potential outcomes reflects both sources of uncertainty. Econometrics refers to this problem as "errors in variables," but it takes on a special meaning when it comes to the interpretation of causal interventions, since it stresses our understanding of what an "observed" potential outcome is.

Practically, this means that we will underestimate our uncertainty about causal effects unless we can account for the uncertainty in our treatment variable on top of our uncertainty in its effect on candidate positioning. All ideal point estimates are uncertainty, but not many studies confront this uncertainty. Since the usefulness of the ideal point model in Chapter 2 is important to explore in this dissertation, it feels appropriate not to ignore posterior uncertainty in its applications. As such, I conduct all sequential-g analyses below using raw posterior samples, rather than posterior means. I describe the exact routine in Section 4.3.2.

4.3 Data and Empirical Approach

I implement the sequential-g routine using candidate data drawn primarily from Bonica (2019), Foster-Molina (2016), and my own estimates.

For the dependent variable, I use Bonica's dynamic Dime scores for House primary candidates in election years 2012, 2014, and 2016. These estimates are generated from a data reduction on campaign contributions data, assuming that campaign contributors give money to ideologically proximate candidates.

For the mediator, I use the Republican share of the presidential vote in the prior election, available in the Bonica data for each candidate running for House. This serves as a broad measure of district voting that closely reflects the partisan composition of the district.

For my independent variable, I use my own district-party ideology estimates generated for the post-2012 districting cycle, estimated from ANES and CCES data for years 2012 through 2018.[4]

Control variables are measured primarily at the congressional district level and are drawn primarily from Foster-Molina (2016). These covariates include racial, educational, religious, and economic features of districts, which I revisit below. Currently I do not have data on primary institutions included in the routine, which limits what I can say about effect modification in primaries vs. caucuses or as a function of primary openness. This is an important part of the story to include in the future.

4.3.1 Model

As mentioned above, I estimate average controlled direct effects separately for incumbents, challengers, and open seat candidates in two parties, totaling six routines altogether. I write a general model that applies regardless of which subset of data is being used. Because the estimation in every data subset proceeds in two stages, I use subscript some parameters 1 and 2 to indicate that they are not fixed across model stages.

The first stage is a mediator-outcome model, predicting the Dime score of candidate i in group g, where a group is a combination of district d and party (p, which is fixed within models for now). Covariate specifications are intended to identify the effect of the previous Republican presidential vote share in district d ($pvote_g$). This is done with the following regression model:

$$\text{Dime}_i = \mu_1 + \tau_1 \bar{\theta}_{g[i]} + \delta \text{pvote}_{d[i]} + \mathbf{x}_{d[i]}^\top \beta_1 + \mathbf{z}_i^\top \gamma + \alpha_{1d} + \varepsilon_i \tag{4.4}$$

where μ_1 is a constant, τ_1 is the coefficient for district-party ideology $\bar{\theta}_g$, δ is the coefficient for the past district vote share. The $\mathbf{x}_d^\top \beta_1$ and $\mathbf{z}_{ig}^\top \gamma$ terms are covariate adjustments for pre-treatment confounders $\mathbf{x}$ and intermediate confounders $\mathbf{z}$. Pre-treatment confounders consist of district-level demographic indicators for the racial composition, college graduation rate, median income, inequality (Gini),

[4] One feature in this design that needs improvement is that it would be nice to have ideology estimates that evolve more over time, so as not to treat district-party ideology as fixed for an entire redistricting cycle, which is the current setup of the model.

unemployment rate, foreign-born population, and evangelical population, and $\mathbf{z}_{ig}$ currently contains an order-3 polynomial function of candidate i's total campaign receipts.[5]

Because many districts contain multiple candidates, I include a district-level intercept α_{1d} that is modeled as a hierarchical Normal error term. Lastly, ε_i is assumed to be Normal.

The second stage of sequential-g estimation requires a demediated outcome. This requires that we isolate the mediator's effect on the outcome in Equation (4.4) and fix it such a way that the mediator falls out of the model.[6] The specification of the first stage implies a demediation function of the following form:

$$\psi_d\left(\text{pvote}'\right) = \hat{\delta}\left(\text{pvote}_d - \text{pvote}'_d\right) \tag{4.5}$$

where *pvote*$'$ is the value of the mediator where the CDE is fixed for all units, and $\hat{\delta}$ is the mediator effect estimate. In the case where the mediator is fixed at zero, this expression reduces to $\hat{\delta} \times pvote_d$. The definition of $\psi_d(\cdot)$ can be more complex if the mediator's effect on the outcome is modeled with interaction terms. This function is subtracted from the outcome data to calculate the demediated outcome, $b\left(\text{Dime}\right)_i$.

$$b\left(\text{Dime}\right)_i = \text{Dime}_i - \psi_d\left(\text{pvote}'\right) \tag{4.6}$$

Because the previous presidential vote share varies from district to district, the demediation function also varies across districts. However, because the mediator is fixed for a given district, all candidates from that district are demediated with the same blipdown function, regardless of their original Dime score.

The demediated outcome is then used in the second stage model, which has a similar specification to the first stage except that the mediator and intermediate confounders are omitted.

$$b\left(\text{Dime}\right)_i = \mu_2 + \tau_2\bar{\theta}_g + \mathbf{x}_{ig}^{\top}\beta_2 + \alpha_{2d} + u_i \tag{4.7}$$

Under sequential unconfoundedness, τ_2 represents the controlled direct effect of district-party preferences on Dime scores. This regression includes residual error term u_i and district error α_{2d}.

[5] Note: this isn't a good choice for intermediate confounding, and it should change. Ask me why!

[6] Or "blipped" out of the outcome value, hence the name "blipdown function."

These models contain hierarchical variance terms, so they are fit with MLE algorithms from the `lme4` package for R.

4.3.2 Algorithm

Because the demediated outcome value in stage two depends on parameter estimates from stage one, our average controlled direct estimate τ_2 should reflect parameter uncertainty from both stages. Acharya, Blackwell, and Sen (2016) derive an analytical variance estimator that deals with this, but unfortunately this estimator does not incorporate the additional uncertainty introduced by the fact that the key independent variable, district-party ideology, is an estimate from a separate measurement model. To overcome this, I implement a simulation routine that propagates model uncertainty from the ideal point estimate into both stages, as well as stage one uncertainty into the second stage.

Suppose that we had a matrix of ideal point samples Θ that had one row for every district-party group and one column for every Markov chain Monte Carlo iteration from the estimation in Chapter 2. We begin by drawing a set of m columns from this matrix, and then for each column m we perform the following routine:[7]

- Estimate the mediator model using the sampled ideal point.
- To carry uncertainty forward from the mediator model, sample 1 coefficient draw from the implied posterior distribution for the mediator effect. This results in a random draw of the mediator effect that combines the uncertainty in the district-party ideal points with the uncertainty in the mediator's effect on the outcome.
- Use each sampled mediator effect to demediate the dependent variable. The demediated value reflects uncertainty in the mediator effect and in the initial ideal point draw.
- Use the demediated outcome to estimate the average controlled direct effect of district-party ideology. Sample one ACDE draw from the implied posterior and store it.

[7] The index m represents one MCMC iteration, with all parameters from that iteration sampled simultaneously. This deals with the fact that ideal point parameters will be correlated within iteration, so they should be supplied to the sequential-g algorithm together.

This routine was performed for $m = 1,000$ ideal point draws, so the result is a distribution of 1,000 ACDE simulations that reflect measurement uncertainty in district-party ideology as well as modeling uncertainty in both stages of sequential-g.

4.4 Findings

I'm running short on time to write these results up gracefully, but here's what we're seeing so far!

4.4.1 Stage 1

Stage 1 estimates the previous presidential vote's effect on candidate ideal points (the mediator-outcome model). Mediator effects in Figure 4.4 show noisy estimates. A one-unit represents a 100 percent change in Republican vote, whereas most Dime scores for candidates live in the [-2, 2] range. So this regression is getting hardly any signal from the presidential vote.

This maybe isn't so crazy, since most of the votes vs. ideal point relationship is between rather than within parties. This echoes back to a key finding in (McCarty, Poole, and Rosenthal 2009), that the within-party relationship between district vote shares and NOMINATE scores was actually quite weak! It appears that these data find an even weaker relationship between district votes and Dime scores.

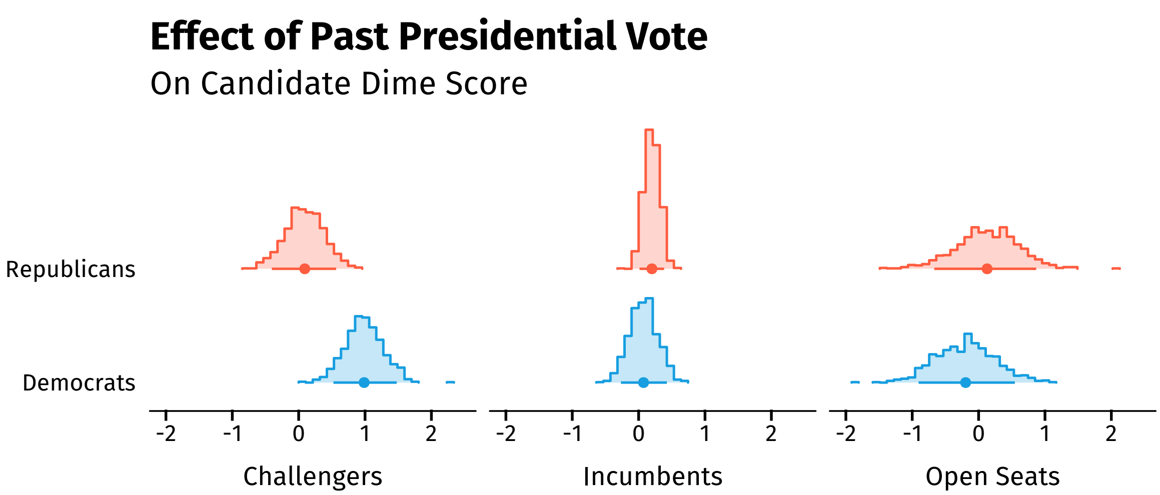

Figure 4.4: Mediator findings

4.4.2 Stage 2

We're seeing average controlled direct effects that are statistically noteworthy in Figure 4.5. Error bars represent 90 percent intervals,[8] which means that most of our simulations are showing ACDEs that have 95 percent posterior probability above zero. I know that I need to do some rescaling of variables to make these effects more interpretable, and once I do that I will be able to put regularizing priors on these effects so as to guard against overfitting in the inference of the treatment effect.

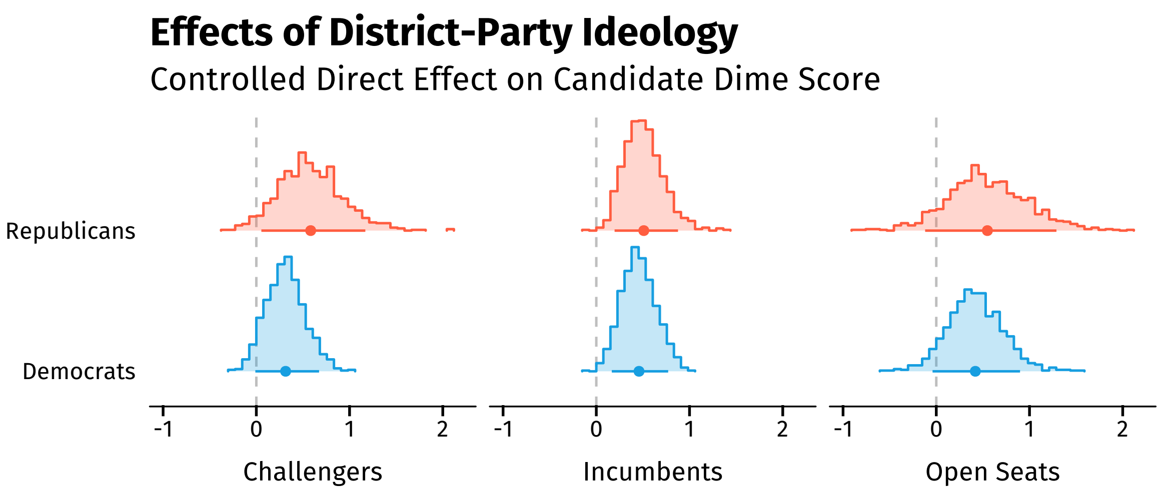

Figure 4.5: ACDE findings

[8] 90 percent intervals are chosen because simulation-based statistics they are more stable estimators of the interval bounds than 95 percent intervals, especially when we don't have a huge number of iterations to work with.

— 5 —

The Constituency Decides: Preferences over Ideological Alternatives

Dropped for APW

References

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.

Ansolabehere, Stephen, James M Snyder, and Charles Stewart. 2001. "Candidate positioning in u.s. House elections." *American Journal of Political Science*: 136–159.

Boatright, Robert G. 2013. *Getting primaried: The changing politics of congressional primary challenges*. University of Michigan Press.

Bonica, Adam. 2019. "Database on ideology, money in politics, and elections: Public version 1.0."

Bonica, Adam. 2014. "Mapping the ideological marketplace." *American Journal of Political Science* 58(2): 367–386.

Burden, Barry C. 2004. "Candidate positioning in u.s. Congressional elections." *British Journal of Political Science* 34(02): 211–227.

Caughey, Devin, and Christopher Warshaw. 2015. "Dynamic estimation of latent opinion using a hierarchical group-level irt model." *Political Analysis* 23(2): 197–211.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. "The statistical analysis of roll call data." *American Political Science Review* 98(02): 355–370.

Downs, Anthony. 1957. *An economic theory of democracy*. New York: Harper; Row.

Foster-Molina, Ella. 2016. "Legislative and district data 1972–2013." https://doi.org/10.7910/DVN/26448.

Fox, Jean-Paul. 2010. *Bayesian item response modeling: Theory and applications*. Springer Science & Business Media.

Greenland, Sander, Judea Pearl, and James M Robins. 1999. "Causal diagrams for epidemiologic research." *Epidemiology*: 37–48.

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, Seth J. 2015. "Institution of nomination and the policy ideology of primary electorates." *Quarterly Journal of Political Science* 10(4): 461–487.

Hirano, Shigeo et al. 2010. "Primary elections and party polarization." *Quarterly Journal of Political Science* 5: 169–191.

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.

Londregan, John. 1999. "Estimating legislators' preferred points." *Political Analysis* 8(1): 35–56.

McCarty, Nolan, Keith T Poole, and Howard Rosenthal. 2009. "Does gerrymandering cause polarization?" *American Journal of Political Science* 53(3): 666–680.

McGhee, Eric et al. 2014. "A primary cause of partisanship? Nomination systems and legislator ideology." *American Journal of Political Science* 58(2): 337–351.

Montgomery, Jacob M, Brendan Nyhan, and Michelle Torres. 2018. "How conditioning on post-treatment variables can ruin your experiment and what to do about it." *American Journal of Political Science* 62(3): 760–775.

Robins, James M, and Sander Greenland. 1994. "Adjusting for differential rates of prophylaxis therapy for pcp in high-versus low-dose azt treatment arms in an aids randomized trial." *Journal of the American Statistical Association* 89(427): 737–749.

Rubin, Donald B. 1978. "Bayesian inference for causal effects: The role of randomization." *The Annals of statistics*: 34–58.

Vansteelandt, Stijn. 2009. "Estimating direct effects in cohort and case–control studies." *Epidemiology*: 851–860.