Do Primaries Work?

Candidate Positioning and Representation in Nominating Elections

Michael G. DeCrescenzo

Contents

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

List of Tables

List of Figures

Acknowledgments

Thanks to [insert people]:

- My social environment (smh)
- Family
- Professors
- Friends
- The Haters
- Strauss
- Cohort of Strauss's class
- Committee
- Warshaw; Treul; Kernell
- Beasts

Chapter One

Primary Elections and Ideological Choice

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 identified a number of positive representational consequences to polarization. Compared to the parties of the early- and mid-1900s that political scientists believed were alarmingly undifferentiated (American Political Science Association 1950), in recent decades the Democratic and Republican Parties have taken increasingly divergent and oppositional stances across a wider variety of issues. Voters have "sorted" into partisan groups that occupy distinct ideological locations in American politics, leading to greater consistency in their issue beliefs, greater ideological abstraction in voter attitudes, and increased political participation (Abramowitz and Saunders 1998; Fiorina, Abrams, and Pope 2005; Layman and Carsey 2002; Levendusky 2009).

Even as polarization has strengthened many aspects of inter-party representation, it affects within-party representation in a number of troubling ways. Party labels provide starker informational and identity cues about candidates than in decades past, but for the typical voter, there is not much of a decision to be made. 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 has increased, there is a sense in which polarization exacerbates the notion of partisan electoral "capture." For most voters, their choices are locked in long before Election Day. Candidates from their preferred parties have been selected through a nomination process, and voters are more likely to abstain from voting when faced with an undesirable candidate than they are to vote against their parties (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 teamsmanship (Rogowski 2016). Voters relax their ideological scrutiny of candidates to cast low-cost votes for their own party, weakening the influence of policy preferences in electoral representation overall.

Hate this

sense in
ich" come on

ore forgiving of
d traits?

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 are relatively weak venue for democratic accountability, does the U.S. electoral system incorporate these preferences in other ways?

When the choice before voters in the general election does not present realistic alternatives, political scientists naturally shift their focus to the nomination of partisan candidates. V.O. Key, for example, famously studied one-party rule in the American South, asking whether competition within the Southern Democratic 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 single-party dominance is a serious limitation. Within-party representation presents interesting questions that apply to far wider contexts. 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.

what am I sayin here?

This dissertation is chiefly concerned with the policy preferences of partisan voters and their role in electoral representation through primary politics. 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), 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 primary representation, 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 preferences on (1) the positioning of candidates and (2) the outcomes of primary elections.

punch up or jus preview?

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 attentive to politics and hold stronger, non-centrist issue preferences compared to the average general-election voter.[1] As a result, competition in the primary stage may present a more acute electoral threat to partisan candidates than general elections do—a "strategic-positioning dilemma" that leads candidates to take ideological issue stances in favor of convergent stances that target the median voter (Aldrich 1983; Brady, Han, and Pope 2007; Burden 2004; Hill 2015).

nsidered cites in
mment

I argue that if we were to construct ideal tests for this theory, the preferences of partisan primary voters play a key role. First, the theory predicts that candidates strategically position themselves to appeal to primary voters. That is, if we could construct an ideological summary of the partisan primary voters in a district, we ought to find that the level of voter conservatism affects the conservatism of the primary candidates' campaign stances. And second, if primary voters present a credible threat to primary candidates (specifically via their policy preferences), they should vote for the candidate who best represents their policy views. We should observe

[1]Primary 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 moderate the strong partisan composition of primary voters (Hill 2015).

that as the degree of primary voter conservatism increases, the probability that a more conservative candidate is nominated increases as well. These two predictions are the core empirical implications of the "strategic positioning dilemma" theory of representation in primaries, and they each require researchers to know the preferences of partisan voters within an electoral district. Yet, the preferences of the partisan constituency are either absent or dangerously misconstrued in the existing literature, and we have been unable to answer these key questions as a result.

1.1 Limitations for estimating ideal points from votes

This section shows the limitations of inferring voter preferences from vote shares, using notation similar to Kernell (2009).

Georgia Kernell (2009) demonstrates the difficulty of inferring district-level preferences (median voter locations) from observed votes. First, she shows that the ordering of district medians cannot be inferred from the vote shares of one election. Suppose we wish to place districts 1 and 2 on an ideological dimension. We observe that the Republican share of the two party vote in each district is p_1 and p_2. Assuming that voter preferences are normally distributed around the median voter (equivalent to the mean voter), then vote shares can be understood as the result of the normal CDF by comparing the candidate's ideal point to the distribution of voter preferences.

$$p_i = \Phi\left(\frac{c - \mu_i}{\sigma_i}\right) \tag{1.1}$$

By inverting the normal CDF, she shows that the difference in medians μ_1 and μ_2 is not proportional to vote shares in each district, but to the z-score of the vote in each district.

$$\mu_2 - \mu_1 \propto Z\left(p_2\right) - Z\left(p_1\right), \tag{1.2}$$

The nonlinear relationship between μ_i and p_i suggests that district preferences are unidentifiable without using multiple elections resulting from the same fixed set of voter preferences μ_i.

Following a similar setup, I demonstrate that the task of inferring the positions of multiple parties is even more difficult. First, rather than assuming that voter preferences in a district are normally distributed with one mean and one dispersion term, we assume that voter preferences are mixture of two distributions (indexed 1 and 2). Each party votes for a Republican candidate akin to the normal CDF as before, but the vote share p for the Republican reflects the size of each partisan constituency in the district as well. Suppose that the size and ideal point of party unaffiliated is represented by a normally distributed error term ε:

$$p = \Phi\left(\frac{c - \mu_1}{\sigma_1}\right)\pi_1 + \Phi\left(\frac{c - \mu_2}{\sigma_2}\right)\pi_2 + \varepsilon, \tag{1.3}$$

where p_i reflects the proportion of the electorate that identifies with party i. Isolating μ_1 and μ_2 in this setup is rather difficult, and the choice of simplifying assumptions is limited. We could assume that the variance of preferences in each party is equal, $\sigma_1 = \sigma_2 = \sigma$, and manipulate the equation somewhat…

$$Z(p)\,\sigma = (c - \mu_1)\,\pi_1 + (c - \mu_2)\,\pi_2, \tag{1.4}$$

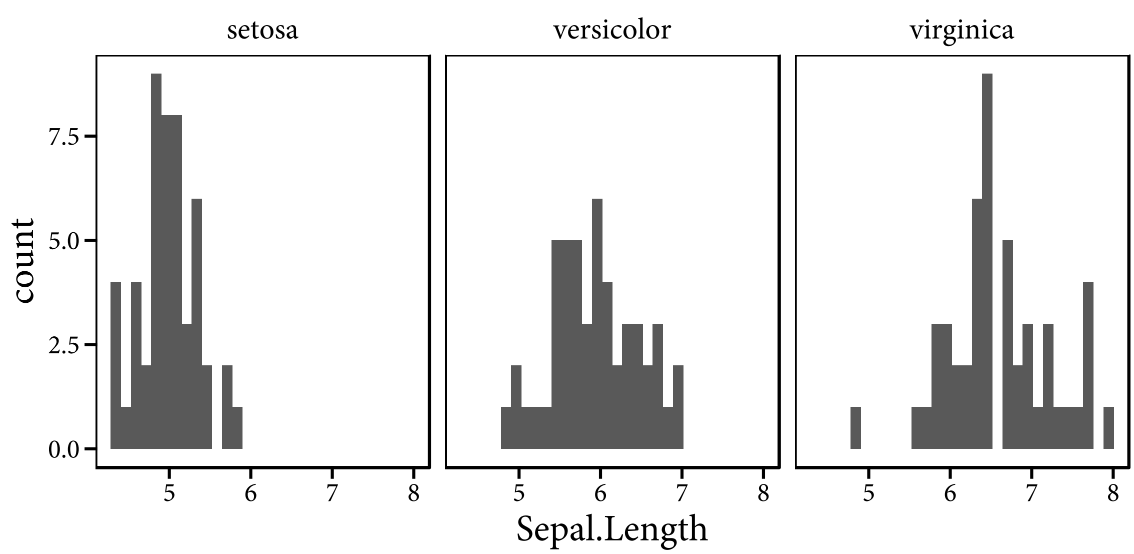
but the limitations are still significant. We could not make a simplifying assumption that $\pi_1 = \pi_2$ for but a handful of districts. We might estimate π_i from survey data

Chapter Two

Modeling the Constituency's Policy Preferences

2.1 Ideal Points and Spatial Models in Political Science

What is the formal model intuition?



2.2 Item Response Theory

2.2.1 Latent variable intuition

Ideology itself is unobservable. It affects the way people feel about policy, but we don't directly observe it. Explicit measures of ideological self identification are not themselves reliable because people don't understand those items and because they aren't great at predicting policy views either

2.2.2 Ability-only model

We begin with a simple model where an individual i gives a conservative response to item j if they derive more utility than they would from a liberal response. Their utility function is made up of their ideal point θ_i and idiosyncratic error ε_{ij}. We can assume that the threshold between liberal and conservative responses is at 0; this is merely a scale and location restriction on what is otherwise an unstructured ideological space.

$$y_{ij} = I\left(\theta_i + \varepsilon_{ij} > 0\right) \tag{2.1}$$

This means $y_{ij} = 1$ when the utility is positive, and $y_{ij} = 0$ otherwise.

Assuming that the idiosyncratic utility is $\varepsilon_{ij} \sim \text{Normal}(0, 1)$ gives us a probit model,

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

where $\Phi(\cdot)$ is the Gaussian cumulative distribution function. This assumption restricts θ such that the probability of a correct response is 50% when $\theta = 0$. The scale of θ is also restricted by the assumption that $\text{Var}\left[\varepsilon_{ij}\right] = 1$.

2.2.3 One-parameter Rasch model

Ability model assumes no "item effects"—there is no systematic variation at the item level that would lean to a correlated effect for one item across individuals.

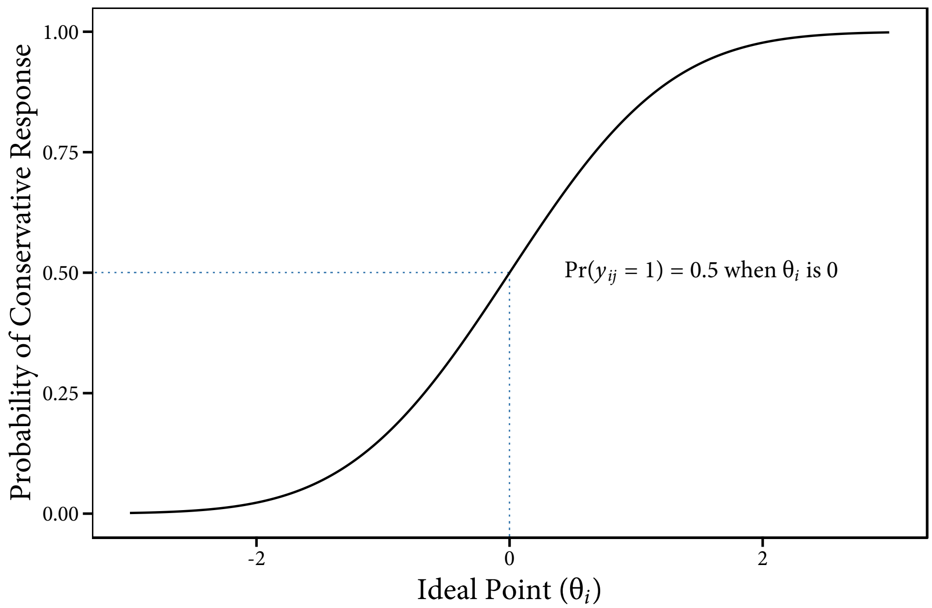


Figure 21: Ability-only model of item response

This is the justification for the one-parameter Rasch model, developed in an educational testing framework. The model defines the probability that individual i answers item j "correctly" as a comparison between an individual i's latent "ability" (θ_i) and a test question j's "difficulty" (κ_j). Assuming that other disturbances affecting item responses (ε_{ij}) are normally distributed, individual i's response to item j is

$$y_{ij} = I\left(\left[\theta_i - \kappa_j\right] + \varepsilon_{ij} > 0\right), \tag{2.3}$$

which yields a probit model for a discrete outcome.

$$\Pr\left(y_{ij} = 1\right) = \Phi\left(\theta_i - \kappa_j\right) \tag{2.4}$$

Adapting this model to the context of survey response, an individual respondent i is asked to choose between two alternatives on policy item j. We let θ_i be i's ideal point on a liberal-conservative dimension, where larger values of θ are more conservative, and κ_j is the midpoint between the liberal and conservative policy alternatives. Respondent i is more likely to give a conservative response on policy j if their ideal point is to the "right" of the item midpoint (their ideal point is nearer to the conservative policy choice), but their response is affected by other normally distributed factors and is thus predicted only probabilistically.

2.2.4 Two-parameter Rasch Model

The item-response model makes sense only under the assumption that policy issues are choices along a single underlying dimension. The one parameter model allows non-ideological factors to affect item responses but assumes that these extraneous factors are independent across items and individuals. Policy issues don't all have the same salience to partisan and ideological debate, however, so political scientists often model policy item responses using a two-parameter model. The two-parameter model includes the additional item parameter β_j, which relaxes the assumption that all items are equally related to latent ideology. This "discrimination" parameter captures how strongly the item divides the responses of liberals and conservatives.

$$\Pr\left(y_{ij} = 1\right) = \Phi\left(\beta_j\left[\theta_i - \kappa_j\right]\right), \tag{2.5}$$

2.2.5 Applications of the IRT Approach

2.3 A Heteroskedastic Model of Party-Public Ideology

This section outlines my ideal point model for local partisan groups. First, I lay out the intuition and notation of a basic individual-level IRT model and its connection to the group-level model. I then lay out my hierarchical model of party-public means and its dynamic form. Finally, I discuss model identification and prior distributions.

2.3.1 Group-Level IRT Setup

To motivate the intuition of the group model, we begin the intuition of the individual and then reparameterize it.

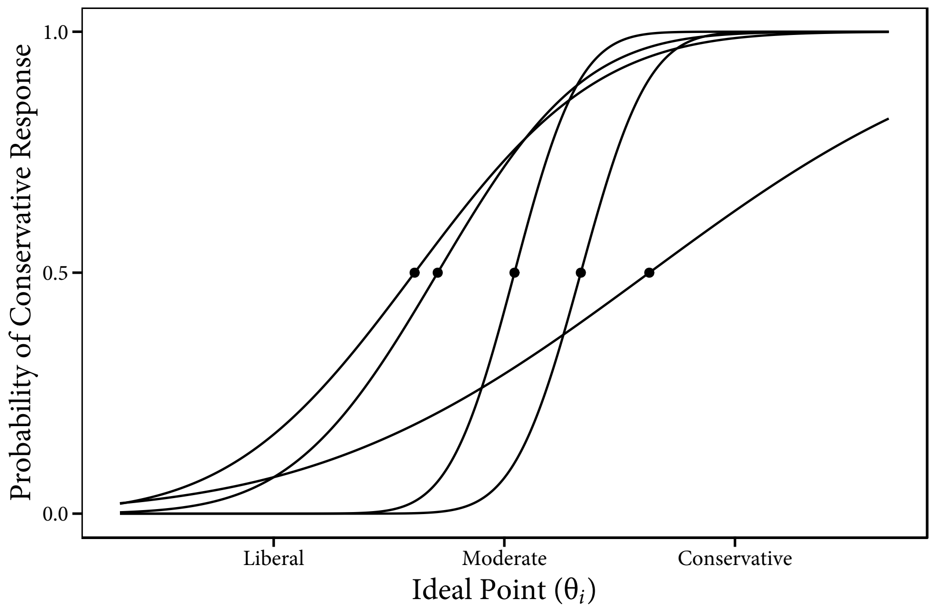


Figure 22: Examples of item characteristic curves for 5 simulated items.

Assume that individuals answer policy items according to their own individual ideal points. We observe a binary response from individual i to item j, which we model as a probabilistic outcome with probability π_{ij}.

$$y_{ij} \sim \text{Bernoulli}(\pi_{ij}) \tag{2.6}$$

Our model for the response probability is given by a probit model, following the utility intuition described above.

$$\pi_{ij} = \Phi\left(\beta_j \left[\theta_i - \kappa_j\right]\right) \tag{2.7}$$

Let $\sigma_j = \beta_j^{-1}$, so item discrimination is instead expressed as a "dispersion" parameter (Caughey and Warshaw 2015; Fox 2010).

$$\pi_{ij} = \Phi\left(\frac{\theta_i - \kappa_j}{\sigma_j}\right) \tag{2.8}$$

Assume that individuals are normal draws from the mean of their group, where a group is a party in a district.[1]

$$\theta_i \sim \text{Normal}\left(\theta_{g[i]}, \sigma_{g[i]}\right) \tag{2.9}$$

The outcome data at the group level, rather than individual Bernoulli outcomes, are expressed as grouped Binomial outcomes Y_{gj}: the number of conservative responses in group g to item j given the total number of responses per group per item, n_{gj}. The probability that a randomly selected individual gives a conservative response (the "true conservative probability" to item j in group g) is $\bar{\pi}_{gj}$.

$$Y_{gj} = \text{Binomial}\left(\bar{\pi}_{gj}, n_{gj}\right) \tag{2.10}$$

The probability for each item-group is given a probit model from the spatial utility model.

$$\bar{\pi}_{gj} = \Phi\left(\frac{\theta_g - \kappa_j}{\sqrt{\sigma_g^2 + \sigma_j^2}}\right), \tag{2.11}$$

where σ_g is the standard deviation of ideal points within group g, which is introduced because we are now estimating the mean response within a group rather than an individual item response. Larger values of σ_g indicate more uncertainty over the item response and attenuate $\bar{\pi}_{gj}$ toward 50%.[2]

2.3.2 Hierarchical Model for Group Means and Scales

les

Estimates for θ_g and scales σ_g are improved using a hierarchical model.

- Using geographic data to improve estimation

---

[1] 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.

[2] The binomial setup assumes that each of the n_{gj} trials is independent conditional on the mean ideal point θ_g and the item parameters κ_j and σ_j. This assumption is violated if individuals in a group answer multiple items—errors are not independent across items. I describe a design-weighting correction for this assumption in Section 2.4.

- Hierarchical model accounts for multiple sources of variation
- Partial pooling estimates for groups without as much data
- First, using traditional notation. Reparameterization in Section 2.4 greatly improves the estimation in Stan.

We consider it a draw a distribution with hypermean $\bar{\theta}_g$ and scalar standard deviation ψ_θ

$$\theta_g \sim \left(\bar{\theta}_g, \psi_\theta\right) \tag{2.12}$$

The hierarchical setup improves estimates by casting the hypermean as a conditional mean from a regression on group data. We specify this regression using geographic-level data from the districts and states where each group is located:

$$\bar{\theta}_g = \beta_{0p[g]} + X_{d[g]}\beta_{p[g]} + \alpha_{s[g]p[g]}, \tag{2.13}$$

where β_0 is a constant, X_d represents district-level data with coefficients β, and α_s represents state effects. A key feature of the hierarchical model is that the parameters are indexed by p and thus dependent on the party to which g belongs. This means there are two constants β_{0p} where $p \in \{1, 2\}$ indexes party. Group-level covariates X_g have coefficient vectors β_p that vary by party as well. We include this flexibility because geographic covariates (such as racial composition) may have different correlations with ideal points depending on the party. This is a departure from the model laid out by Caughey and Warshaw (2015), which holds the geographic regression fixed for all groups in the data.

We use a similar hierarchical regression for group scales (suppressing the g subscript which is implied by the combination of d and p).

$$\log\left(\sigma_g\right) \sim \text{Normal}\left(\delta_{0p} + X_d\delta_p + \eta_{sp}, \psi_\sigma\right), \tag{2.14}$$

where δ_{0p} represents constants for each party, δ_p is a party-specific vector of coefficients on district features, and η_{sp} are party-specific state effects.

The state effects are in turn regressions on state features.

$$\alpha_{sp} \sim \text{Normal}\left(Z_s\gamma_p + \rho_{r[s]p}, \psi_\alpha\right), \tag{2.15}$$

$$\eta_{sp} \sim \text{Normal}\left(Z_s\zeta_p + \xi_{r[s]p}, \psi_\eta\right), \tag{2.16}$$

where Z_s contains state covariates which have party-specific coefficients γ_p (for group means) or ζ_p (for group scales). Each state effect is a function of a party-specific region effect ρ_{rp} (for group means) and ξ_{rp} (for group scales) for Census regions indexed r.

2.3.3 Dynamic Model

2.3.4 Priors and Identification

2.4 Implementation in Stan

ly caveats about
erences to Stan

Hierarchical models often have posterior distributions whose curvature presents difficulties for sampling algorithms (Betancourt and Girolami 2015; Papaspiliopoulos, Roberts, and Sköld 2007). To improve the estimation in Stan, I parameterize the hierarchical models in the "non-centered" rather than the "centered" form. Whereas the centered form considers θ_g as a random draw from a distribution, the non-centered parameterization considers θ_g as a function of the hypermean and a random error.

$$\theta_g = \beta_{0p} + X_g\beta_p + \alpha_{s[g]} + z_g\varepsilon \tag{2.17}$$

where $z_g\varepsilon$ represents a group-level error term. It is composed of a z-score that is Normal$(0, 1)$ and a scale parameter ε. The non-centered paramaterization has the same algebraic behavior as the centered parameterization, but it has the practical effect of improving Monte Carlo sampling by de-correlating the parameters in the hierarchical model. We can also de-center the $\alpha_{s[g]}$ term

$$\alpha_{s[g]p} = Z_{s[g]}\gamma_p + u_{sp}\tau \tag{2.18}$$

where u_{sp} is a z-score distributed Normal $(0, 1)$, and τ is a scale factor.

We also have a hierarchical model that predicts the ideal point standard deviation within each group, σ_g. This makes the model "heteroskedastic"—we are modeling the mean ideal point within each group and the variance, conditional on hierarchical covariates. The model for σ_g in non-centered form is as follows:

$$\log(\sigma_g) = X_g\delta_p + Z_{s[g]}\eta_p + m_g\nu + m_{sp}\lambda \tag{2.19}$$

Where X_g and Z_{sp} are the same group- and state-level covariates as the above regression, δ_p and η_p are party-varying coefficients. The terms $m_g\nu$ and $m_{sp}\lambda$ are "factored" error terms for groups and states, where m_g and m_{sp} are each distributed Normal $(0, 1)$, while ν and λ are scale factors.

Factoring

IRT Model

Noncentered Hierarchical Model

Any other vector/matrix tricks?

Read file into an appendix?

2.5 Simulation and Model Validation

Doing a non-centered parameterization

Chapter Three

Candidate Positioning in Primary Elections

U.S. House primary data from Pettigrew, Owen, and Wanless (2016).

Chapter Four

The Constituency Decides: Preferences over Ideological Alternatives

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Appendix A

Group IRT Model

Appendix B

Colophon

This document was produced using `R`, R Markdown, LaTeX. The document was built using the `bookdown` package for `R`. The document template is a variation on TJ Mahr's `buckydown` template, which was itself an adaptation of documents designed by and for students at the Universities of Washington and Wisconsin. The PDF is typeset using pdfTeX. The body text is *MinionPro-LF* in 12pt size.

The data and source code for this dissertation have been organized into an online Git repository on Bitbucket. A hard copy of the thesis can be found in the University of Wisconsin library system.

- Git repository: https://bitbucket.org/mikedecrescenzo/dissertation
- `huskydown`: https://github.com/benmarwick/huskydown
- `bookdown`: https://github.com/rstudio/bookdown

This version of the thesis was generated on 2019-05-22 09:28:19. The repository is currently at this commit:

```
## Commit:  3fa6d70d798e2cbed6e82912d1cb34c618518c9f
## Author:  Michael DeCrescenzo <mgdecrescenzo@gmail.com>
## When:    2019-05-14 12:56:33
##
```

```
##      wow trying again, actually homoskedastic
##
##
## 1 file changed, 2 insertions, 2 deletions
## code/dgirt/stan/long/long-homo-mlm.stan | -2 +2  in 2 hunks
```

References

Abramowitz, Alan I, and Kyle L Saunders. 1998. "Ideological realignment in the us electorate." *The Journal of Politics* 60(03): 634–652.

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.

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

Bartels, Larry M. 2000. "Partisanship and voting behavior, 1952-1996." *American Journal of Political Science*: 35–50.

Betancourt, Michael, and Mark Girolami. 2015. "Hamiltonian monte carlo for hierarchical models." *Current trends in Bayesian methodology with applications* 79: 30.

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.

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

Campbell, Angus et al. 1960. New Yrok: John Wiley and Sons 77 *The american voter*.

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

Cohen, Marty et al. 2009. *The party decides: Presidential nominations before and after reform*. University of Chicago Press.

Downs, Anthony. 1957. "An economic theory of democracy."

Fiorina, Morris P, Samuel J Abrams, and Jeremy C Pope. 2005. *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.

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

Geer, John G. 1988. "Assessing the representativeness of electorates in presidential primaries." *American Journal of Political Science*: 929–945.

Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan hearts and minds*. New Haven, CT: Yale University Press.

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.

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.

Kernell, Georgia. 2009. "Giving order to districts: Estimating voter distributions with national election returns." *Political Analysis* 17(3): 215–235.

Key, V.O. Jr. 1949. "Southern politics in state and nation."

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).

Levendusky, Matthew. 2009. *The partisan sort: How liberals became democrats and conservatives became republicans*. University of Chicago Press.

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

Norrander, Barbara. 1989. "Ideological representativeness of presidential primary voters." *American Journal of Political Science*: 570–587.

Papaspiliopoulos, Omiros, Gareth O Roberts, and Martin Sköld. 2007. "A general framework for the parametrization of hierarchical models." *Statistical Science*: 59–73.

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).

Pettigrew, Stephen, Karen Owen, and Emily Wanless. 2016. "U.S. House primary election results (1956-2010)." https://doi.org/10.7910/DVN/26448.

Rahn, Wendy M. 1993. "The role of partisan stereotypes in information processing about political candidates." *American Journal of Political Science*: 472–496.

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.